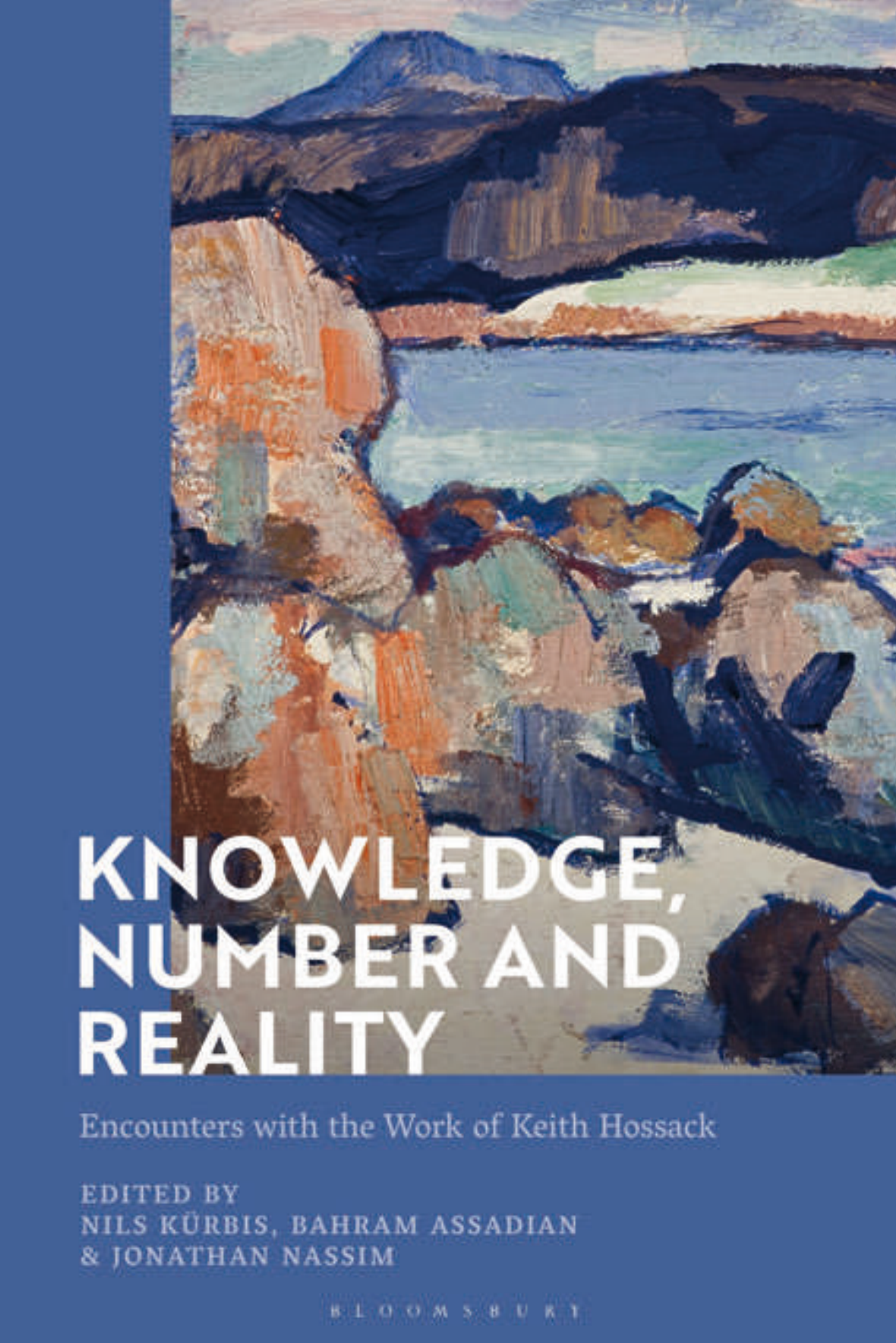




KNOWLEDGE, NUMBER AND REALITY

Encounters with the Work of Keith Hossack

EDITED BY
NILS KÜRBIS, BAHRAM ASSADIAN
& JONATHAN NASSIM

BLOOMSBURY

Knowledge, Number and Reality

Knowledge, Number and Reality

Encounters with the Work of Keith Hossack

Edited by

Nils Kürbis, Bahram Assadian and Jonathan Nassim

BLOOMSBURY ACADEMIC
LONDON • NEW YORK • OXFORD • NEW DELHI • SYDNEY

BLOOMSBURY ACADEMIC
Bloomsbury Publishing Plc
50 Bedford Square, London, WC1B 3DP, UK
1385 Broadway, New York, NY 10018, USA
29 Earlsfort Terrace, Dublin 2, Ireland

BLOOMSBURY, BLOOMSBURY ACADEMIC and the Diana logo are trademarks of
Bloomsbury Publishing Plc

First published in Great Britain 2022

Copyright © Nils Kürbis, Bahram Assadian, Jonathan Nassim and Contributors 2022

Nils Kürbis, Bahram Assadian and Jonathan Nassim have asserted their right under the
Copyright, Designs and Patents Act, 1988, to be identified as Editors of this work.

Cover image: Iona Landscape: Rocks, by Samuel John Peploe, 1925-1927. Oil on canvas.
(Photo by National Galleries Of Scotland/Getty Images)

All rights reserved. No part of this publication may be reproduced or transmitted
in any form or by any means, electronic or mechanical, including photocopying,
recording, or any information storage or retrieval system, without prior
permission in writing from the publishers.

Bloomsbury Publishing Plc does not have any control over, or responsibility for,
any third-party websites referred to or in this book. All internet addresses given
in this book were correct at the time of going to press. The author and publisher
regret any inconvenience caused if addresses have changed or sites have ceased
to exist, but can accept no responsibility for any such changes.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

ISBN: HB: 978-1-3501-8643-9
ePDF: 978-1-3501-8644-6
eBook: 978-1-3501-8645-3

Typeset by RefineCatch Limited, Bungay, Suffolk

To find out more about our authors and books visit www.bloomsbury.com
and sign up for our newsletters.

Contents

Notes on Contributors	vi
Introduction <i>Nils Kürbis, Bahram Assadian and Jonathan Nassim</i>	1
1 A Summary of My Current Views <i>Keith Hossack</i>	31
2 Confronting Facts: On Hossack's <i>The Metaphysics of Knowledge</i> <i>Mark Sainsbury</i>	37
3 Who Knows? <i>Mary Margaret McCabe</i>	49
4 Knowledge-first Epistemology and the Input Problem <i>Scott Sturgeon</i>	69
5 Perceiving X = Consciousness of Perceiving X. Hossack and Brentano on the Identity Thesis <i>Mark Textor</i>	87
6 Facts, Knowledge and Knowledge of Facts <i>Bernhard Weiss</i>	101
7 Necessity, Conditionals and Apriority <i>Keith Hossack</i>	117
8 A Commentary on 'Necessity, Conditionals and Apriority' <i>Dorothy Edgington</i>	131
9 The Mathematicians' Use of Diagrams in Plato <i>Tamsin de Waal</i>	143
10 Generality <i>Nils Kürbis</i>	161
11 We Belong Together: A Plea for Modesty in Modal Plural Logic <i>Simon Hewitt</i>	177
12 Aristotelian Aspirations, Fregean Fears: Hossack on Numbers as Magnitudes <i>Øystein Linnebo</i>	189
13 Mathematical Structures, Universals and Singular Terms <i>Bahram Assadian</i>	203
14 Arithmetic in a Finite World <i>Peter Simons</i>	217
Index	229

Notes on Contributors

Bahram Assadian is a fixed term lecturer in the Department of Philosophy at the University of Amsterdam, and a member of the Institute for Logic, Language and Computation (ILLC). He received his PhD in 2016 from Birkbeck College, University of London, co-supervised by Keith Hossack. He has published several papers on the philosophy of mathematics, the philosophy of language, and metaphysics.

Dorothy Edgington studied at Oxford. Most of her career was spent at Birkbeck College, University of London, interrupted by five years (1996–2001) and then three years (2003–6) as a professor at Oxford, in the latter period holding the Waynflete Chair of Metaphysics. She is now an emeritus professor of both institutions. She is a Fellow of the British Academy. Her main research interests are conditionals, vagueness and other kinds of indeterminacy, applying the notion of probability in these areas.

Simon Hewitt completed his PhD on plural logic at Birkbeck College, University of London, supervised by Keith Hossack. Since then he has taught at the University of Leeds, and takes an interest in the philosophy of religion, as well as metaphysics and the philosophies of logic and mathematics. He is the author of several articles.

Keith Hossack was schooled at Airdrie Academy. At Edinburgh he took a degree in Philosophy, and also a degree in Mathematics and Physics. His Oxford doctorate was under the external supervision of Dorothy Edgington. After spells as a civil servant and a schoolteacher he was for many years a lecturer at King's College London. Currently he is Reader in Philosophy at Birkbeck College, University of London, where he recently became part-retired in order to have more time for research.

Nils Kürbis is associate professor of philosophy in the Department of Logic at the University of Łódź, Alexander von Humboldt Research Fellow at the Department of Philosophy I of the University of Bochum and Honorary Research Associate at the Department of Philosophy of University College London. He obtained his PhD from King's College London studying under Keith Hossack. He has published articles on negation, denial and falsity, the proof theory of definite descriptions and a book, *Proof and Falsity. A Logical Investigation*. His current research aims to formulate a proof-theoretic semantics for modal logic.

Øystein Linnebo is Professor of Philosophy at the University of Oslo, having previously held positions at Bristol, London, and Oxford. His main research interests lie in the philosophies of logic and mathematics, metaphysics, and early analytic philosophy

(especially Frege). He has published more than sixty scientific articles and is the author of three books, most recently *Thin Objects: An Abstractionist Account* (2018) and *The Many and the One* (with Salvatore Florio, 2021).

Mary Margaret McCabe ('MM') works on ancient philosophy, on ethics and on the philosophy of medicine; she is the author of *Plato on Punishment* (1981), *Plato's Individuals* (1994), *Plato and his Predecessors: The Dramatisation of Reason* (2000) and *Platonic Conversations* (2015). She is Professor of Philosophy Emerita at King's College London, a Fellow of the British Academy, Honorary Fellow of Jesus College, Cambridge and a Fellow of King's College London. MM is the Chair of Trustees of the charity Philosophy in Prison, which provides and supports philosophical discussion for prisoners in the UK.

Jonathan Nassim is an independent scholar. He took his doctorate on theories of possibility at Birkbeck College, University of London, under the supervision of Keith Hossack. He has published widely including in the journals *History and Philosophy of Logic* and *Theoria*.

Mark Sainsbury (marksainsbury.net) is Jesse H. Jones Regents Professor in Liberal Arts in the Philosophy Department at the University of Texas at Austin. Before coming to Austin in 2002, he taught in the University of London (Bedford College, King's College), and at the universities of Essex and Oxford. His main philosophical interests are philosophy of language and mind. He is the author of eight books, most recently *Thinking About Things*.

Peter Simons is Professor Emeritus of Philosophy, Trinity College Dublin, Ireland. He studied mathematics and philosophy at the University of Manchester, and previously taught at Bolton, Salzburg and Leeds. The author or co-author of five books and some 300 papers, he specializes in metaphysics (pure and applied), philosophy of logic, and the history of logic and philosophy in Central Europe. He is a member of the British, European, Irish and Polish Academies.

Scott Sturgeon's first job was in Keith Hossack's department at King's College London, where he learned a great deal from Hossack. Sturgeon next moved to Keith's current department at Birkbeck College, University of London, where he learned a great deal from Hossack's supervisor Dorothy Edgington. Then he was Professor of Philosophy at the University of Oxford and Tutorial Fellow at Wadham College, and then Professor of Philosophy at the University of Birmingham.

Mark Textor is Professor of Philosophy at King's College London. He has worked on History of Analytic Philosophy (especially Frege), Austrian Philosophy and Philosophy of Mind and Language. His most recent book is *Brentano's Mind*. His book *The Disappearance of the Soul and the Turn against Metaphysics: Austrian Philosophy 1874–1918* is forthcoming.

Tamsin de Waal is currently a Visiting Scholar in Philosophy at Columbia University. She has previously taught at King's College London and University College London. Her primary research interests are in the areas of Ancient Philosophy, especially Plato's epistemology and metaphysics, and Philosophy of Mathematics.

Bernhard Weiss is Professor of Philosophy at the University of Cape Town. He writes mainly in the areas of the philosophies of language and logic; and has published two books, *Michael Dummett* (2002) and *How to Understand Language* (2010).

Introduction

Nils Kürbis, Bahram Assadian and Jonathan Nassim

Keith Hossack is a realist and a rationalist. His views are bold, controversial, unorthodox, sometimes outrageous, but always forcefully argued. He has published two monographs: *The Metaphysics of Knowledge* (2007) and *Knowledge and the Philosophy of Number: What Numbers Are and How They Are Known* (2020), together with a number of influential papers in philosophical journals. (See Chapter 1 for a full bibliography.) In this volume, we collect together new articles by his students and colleagues, many of whom have become his friends, engaging with all aspects of his work: metaphysics, epistemology, the philosophy of mind, logic and the philosophy of mathematics. Because Hossack's interests are so broad and systematic, this volume will be of interest not only to those who know his work, but to anyone engaging with the central questions in analytic philosophy.

This book grew out of the conference 'Reviving Rationalism. A Celebration of the Work of Keith Hossack', held in his honour on 26 November 2016 at Birkbeck College, University of London. It was organized by two of the editors and Simon Hewitt, one of the contributors to this volume. As Hossack was reducing his teaching duties at Birkbeck College, where he is a Reader in Philosophy, we thought it was a good time to celebrate his contribution to philosophy, as a thinker and teacher – not to mention, as *our* teacher. Thanks to a Mind Association Conference Grant and to Hallvard Lillehammer, who secured a grant from Birkbeck College, we were able to hold the conference. This book is a record of recent engagements with Hossack's work which have grown out of it.

We would like to thank all the contributors who have made this volume possible, and everyone whom we approached but who were prevented from contributing due to other commitments, our time frame or the theme of the volume. Apologies to all of Hossack's friends and colleagues we did not ask. The project proved very popular, and we are aware that many more could have contributed – at fourteen contributions we needed to draw a line. Thank you also to the referees for Bloomsbury Press for their enthusiasm for our project and to Colleen Coalter, Becky Holland and their team for their support during the process of publication.¹

We divide up Hossack's oeuvre, somewhat artificially given the systematicity of this thought, into three themes: Realism, Knowledge and Rationalism. Our introduction will follow this outline. Hossack takes much of his inspiration and starting point from Russell's writings, one of his philosophical heroes. In relation to Realism, we will pay particular attention to how Hossack's views interact with Russell's. Then we'll move on

to explore the interaction between Hossack's work and the contributions to this volume.

The notion of realism central to Hossack's work is the doctrine that there are universals which exist independently of the particulars that instantiate them. Hossack adopts Russell's account of universals (Russell 1912: 145ff), according to which a universal is an aspect of resemblance (Hossack 2007: 34ff). Such an account of what a universal is may be uncontroversial. Philosophical controversy arises over the further question whether there are any universals and, if so, what kind of things they are.

Hossack and Russell observe that some things really resemble each other: it is not just that they *appear* to someone to resemble each other, or that they are *perceived* as similar. Socrates and Plato resemble each other in being mortal. Their mortality is not a question of what someone may consider them to be. They really are mortal, no matter whether someone conceives them as such. If things resemble each other objectively, then, Russell and Hossack conclude, there must be universals. (See Hossack 2007, ch. 2.1, Hossack 2020, ch. 1.) If two things really resemble each other in a certain respect, then they share the universal that is the respect or the way in which they resemble. Socrates and Plato both share the universal *mortality*, or they both instantiate it. Some universals are not instantiated by particulars but by other universals. Red and green resemble each other in being colours, so the universals *redness* and *greenness* instantiate the universal *colour*. The referents of the concepts 'square' and 'round' resemble each other in being shapes, so the universals *squareness* and *roundness* instantiate the universal *shape*.

The existence of universals, in turn, explains the nature of resemblance. Things resemble each other because they literally have something in common: the universal they both share. Maybe everything resembles anything in some aspect or other, or can be conceived to do so, but some things also fail to resemble each other in some respects. Wherever there are things that are not similar, there are aspects in which they do not resemble.² This, too, is explained by universals. If two things do not share a universal, they do not resemble each other in this respect. Although Socrates and Thrasymachus resemble each other in being mortal, they do not resemble each other in being wise. Socrates is wise. Thrasymachus is not wise. This is a matter of fact. But Thrasymachus resembles Meletus in this latter respect. Hossack draws a further conclusion: that there must be a universal *negation* (Hossack 2007: 62ff) that accounts for the failure of Thrasymachus and Meletus to instantiate the universal *wisdom*.

Negative facts, facts involving universals that are not instantiated by particulars or other universals, play a central role in Hossack's metaphysics, not least because the more precise account of the distinction between universals and particulars offered by Hossack in the development of his metaphysics is founded on the distinction between negative and positive facts, and his acceptance of the existence of both kinds. The acceptance of negative facts is repugnant to many metaphysicians. Russell famously observed that 'there is implanted in the human breast an almost unquenchable desire to find some way of avoiding the admission that negative facts are as ultimate as those that are positive' (Russell 1919b: 4). He reports that his view that there are negative facts nearly provoked his students at Harvard to riot: 'The class would not hear of there being negative facts at all' (Russell 1919a: 42). Most metaphysicians believe that everything that exists is essentially positive. Russell, however, at least during the period

of his thinking that led to *The Philosophy of Logical Atomism*, argued that negative facts cannot be avoided, and that any attempt to do so only brings them back in another guise. (See Russell 1919a: 42ff and 1919b: 4ff.)

Hossack, and Russell circa 1918, therefore share the unorthodox view that there are negative facts, but they differ strongly over their characterization. According to Russell, the positive fact that *a* stands in relation *R* to *b* contains the relational universal *R* and the particulars *a* and *b*, while the negative fact that *a* does not stand in relation *S* to *b* also contains no more than the relational universal *S* and the particulars *a* and *b*: 'It must not be supposed that the negative fact contains a constituent corresponding to the word "not." It contains no more constituents than a positive fact of the correlative positive form. The difference between the two forms is ultimate and irreducible' (Russell 1919b: 4). While for Russell, positive and negative facts are not distinguished by their constituents, but only by their form, or the way the constituents are assembled to form the fact, Hossack, by contrast, accepts precisely what Russell denies, and explains negative facts in terms of the universal *negation*: it is the referent of the symbol for negation in logic. According to Hossack, the negative fact that *a* does not stand in relation *S* to *b* contains the universal *negation*, the relational universal *S* and the particulars *a* and *b*.

Negative facts are also important in other areas of Hossack's metaphysics. He likes to point out that omissions have causal powers: you forgot to water the plants and they died. The omission to water the plants is a negative fact: the plants were not watered. This negative fact causes the death of the plants. Likewise, a bridge may collapse because of the negative fact that it did not have enough rivets.

Once an ontology of universals is accepted, other considerations than aspects of resemblance may come into play in tackling the question of which universals there are. It may be that not every resemblance requires its own universal, but it suffices if every aspect of resemblance is explained in terms of some universals. There need not be a universal *non-wisdom*, but it suffices that there are the universals *negation* and *wisdom* to explain the resemblance between Thrasy Machus and Meletus in this respect. It is worth drawing a distinction between properties and universals. Not every property need be a universal: although there is the property *non-wisdom*, there may not be such a universal. Some properties are universals, others are more accurately described as compounds of universals, such as the property *non-wisdom*, which is a compound of the universal *negation* and (the universal or property) *wisdom*.

Some properties are not shared by some objects: one object has a property that the other doesn't. Some properties are even had by nothing at all. Nothing has the property *unicorn*. Worse still, some properties could not be had by anything. Nothing could have the property *non-self-identical*. But they are nonetheless properties, according to Hossack (2007: 58). Hence, according to him, there are actually and even necessarily uninstantiated properties.

Hossack also accepts that universals may exist even if they are uninstantiated. For instance, according to Hossack there are logical and mathematical universals. But nothing in the world, as it is, instantiates the geometrical universal *roundness*. There are things that more or less approach being round in the perfect, geometrical sense, but there is nothing such that every point of its circumference is equidistant to its centre.

Whether there may also be necessarily uninstantiated universals would appear to be a further question.

This goes against a popular view going back to Aristotle that only those properties and universals exist which are instantiated. Another popular view is that properties of spatial objects are located where their instances are. This view, too, is not one Hossack agrees with: if uninstantiated properties exist, they cannot be located where their instances are, as there are none.

Many philosophers agree that there are properties or universals, but not many will follow Hossack to the extreme realism he is prepared to accept, where even *negation* is a universal, and where properties and possibly even universals may be necessarily uninstantiated. Hossack is not afraid of drawing controversial conclusions from apparently self-evident first principles.

As we will see, some passages of *The Metaphysics of Knowledge* may sound as if Hossack tends towards a view according to which universals are abundant: a universal is everything that is not a particular, and thus it would appear that any meaningful predicate corresponds to a universal. But a clear counterexample is Hossack's rejection of the view that the predicate 'true', while meaningful, stands for a universal (Hossack 2007: 88). Truth is defined as correspondence to the facts, but that my belief corresponds to the facts is not a property it has, and there is no universal *truth*. Hossack draws the consequence that there is no such link between true *beliefs* and the facts to which they correspond as there is between mind and world when a fact is *known*.

Elements of a sparse view of universals are more prominent in *The Metaphysics of Knowledge*, and this is how Hossack intends to be understood. According to the sparse view, only some meaningful predicates correspond to universals, namely those that latch on to fundamental features of reality. Hossack accepts Socrates' account of a good definition given by Plato in the *Phaedrus* (265e–266a): to give a good definition is to 'cut up each kind according to its natural joints'. 'Plato's claim is that scientific progress requires the discovery of the principles of classification that divide things according to their real resemblances and real differences. The laws of a science connect the properties that are of classificatory importance; thus the very idea of a law of nature is inseparably connected with real or objective resemblance' (Hossack 2007: 35f). Which universals there are is a question of what the world is ultimately like and what laws govern it. They are discovered by us in our attempts to produce optimal descriptions of the world.

In Hossack's view (2007: ch. 2.1, 2020: ch. 1), realism about universals, contrary to its nominalist rivals, can explain the difference between a natural class and a merely miscellaneous collection. Hossack argues for a realist position about mathematical entities too. In his view, numbers are universals. And precisely because of this, our knowledge of them is no more mysterious than our knowledge of any other universals. In fact, the source of epistemological puzzlements about the existence of numbers lies in the thesis that numbers are abstract *objects*. Hossack argues that this thesis traps us in the sceptical conclusion that we cannot have knowledge of numbers and number-theoretic propositions. The right conclusion to be drawn from this sort of scepticism, Hossack claims, is that numbers are not particulars. They are universals.

Abundant or sparse, Hossack's account of universals locates them in the vicinity of Plato's Realm of the Forms. His realism can be described as a species of Platonism, and

Platonist themes abound in his work. Russell himself describes his theory of universals as ‘largely Plato’s, with merely such modifications as time has shown to be necessary’ (Russell 1912: 142f). In conversation, Hossack likes to emphasize that Plato and Russell wrestled with the same fundamental problems, as does he, following in their footsteps.

Though fundamental to Hossack’s metaphysics, *universal* and *particular* are not its most fundamental theoretical notions. These are the notions of *knowledge* and of *fact*. According to Hossack, knowledge is a metaphysically and conceptually primitive notion. *The Metaphysics of Knowledge* defends this view in great detail and applies it to a variety of philosophical issues in metaphysics, epistemology and the philosophy of mind in a demonstration of its strength and versatility. The notion of fact is required as that which is known, as one of the relata of the primitive notion of knowledge, and as that in which universals and particulars are combined.

A primitive notion is one that cannot be defined or analysed any further. This does not mean that we cannot say anything informative or illuminating about it. For instance, being a primitive notion, knowledge is not a species of belief. In particular, if knowledge is primitive, then it is not justified true belief, as in the definition of knowledge first propounded in Plato’s *Meno* (98a), though already shown to be problematic in the *Theaetetus* (201cff). Hossack subscribes to a causal thesis about the relation between knowledge and belief and, because knowledge is primitive, he rejects the commonly held ‘constitutive theses’. According to Hossack, knowledge is caused by beliefs and the exercise of mental faculties, but it is not constituted by them. In particular, knowledge is not constituted by justified true belief plus something or other. The primitive notion of knowledge Hossack has in mind is a direct relation of awareness between a mind and that which is known. This is a relation between mind and world that is not mediated by ideas or thoughts or mental representations. The mind that knows stands in direct contact with that which is known.

Consequently, the primitive notion of knowledge is not a propositional attitude either, and the notion of knowledge on which Hossack’s view is founded must be distinguished from the complex relation expressed by the sentences that are used in knowledge attributions. According to Hossack, in the attribution ‘X knows that *s*’, the sentence *s* serves to report not only the fact of which X is aware, but also the content of the mental event in virtue of which X apprehends the fact in question. Hossack therefore rejects the theory that the attribution ‘X knows that *s*’ expresses a propositional attitude, for an attitude is a relation between a mind and a mode of presentation, or in alternative terminology, a content, a Fregean Thought, or a proposition. Therefore if the attribution ‘X knows that *s*’ expressed only a propositional attitude, X would stand only in a mediated relation to the fact that *s*: the relation between the mind and the fact would be mediated by the mode of presentation and it would only be to the latter that the mind would stand in a direct relation.³

Here another Hossackian theme emerges: the refutation of scepticism, also a concern he shares with Russell. If knowledge were merely a species of belief, I could be in the same state of mind as someone deceived by an evil demon. If beliefs or ideas or propositions or other modes of presentation mediated between the mind and the world, and knowledge arose when these mediators in turn relate correctly to the world,

i.e. are true and arrived at in suitable fashion, then the possibility contemplated by scepticism may actually be the case. The sceptic posits that I could be in the same epistemic state as someone who is deceived by an evil demon. Hossack responds that I could not be in such a state, as knowledge is not a species of belief, and if I know, then I stand in the knowledge relation to facts in the world, and hence I am not deceived by an evil demon (Hossack 2007: 16). Believing and knowing may be phenomenologically indistinguishable to the mind that believes and knows, but they are distinct nonetheless: one brings the mind into direct contact with the facts, the other does not. Thus this kind of scepticism is built on a false assumption.

That which is known is a fact. Hossack's work relies on a substantial theory of facts. 'Realism, the theory of universals, needs to be completed by the addition of a theory of facts' (Hossack 2007: 45). In explicating the notion of fact, Hossack also characterizes the kind of relation that knowledge is. A fact is typically a combination of universals and particulars, such as that Socrates is wise: the fact combines the universal *wisdom* and the particular Socrates. Furthermore, facts combine universals and particulars in orderly fashion. There is a difference between the fact that Romeo loves Juliet and the fact that Juliet loves Romeo: one could obtain without the other.

The notion of fact is a fundamental metaphysical primitive of Hossack's philosophy, and with the facts comes the combination of universals and particulars into facts. Hossack argues that realism and the theory of facts go hand in hand: a satisfactory theory of universals requires the theory of facts, and a satisfactory theory of facts requires realism (Hossack 2007: 34). The notions of universal and particular are explicated beyond their initial characterization, as aspects of resemblance, in terms of how they combine to form facts.

According to Hossack, 'realism takes the highest categories to be particular and universal: it claims that without exception, everything that exists falls into one category or the other' (Hossack 2007: 34). It might be worth pointing out a consequence. Bradley's Regress is the problem that, if instantiation were a universal, then for *a* to instantiate *B*, there should be a further kind of instantiation which is instantiated by *a*, *B* and the universal *instantiation*. Clearly, the same considerations apply to this further kind of instantiation, and so on *ad infinitum*. And clearly, they also apply to the combination relation. To avoid Bradley's Regress, the combination relation must be something other than a universal. Combination certainly is not a particular. Consequently, the combination relation falls outside the categories of universals and particulars. It is syncategorematic, as the scholastics might say, and cannot properly be said to exist. Facts are the rock bottom of Hossack's metaphysics: in those that exist, universals and particulars are combined. Instantiation of a universal by a particular is explained in terms of facts, not conversely.

Russell concluded from his explanation of the nature of universals and the objective resemblance of things in certain respects that there must be universals. It is a further question whether there are also particulars. Russell admitted the possibility that there are no particulars (Russell 1911: 10ff). To tackle this issue, we need a definition of 'particular'. So far, we have a heuristic account of the nature of universals, but a sharper definition of this notion is also desirable and available on the basis of the definition of 'particular'. Hossack's definition of 'particular' and 'universal' relies on the existence of

negative facts, facts that contain the universal *negation* in a prominent position. The latter is the referent of the symbol for 'not' in formal logic. It works like any other universal, in that the combination relation forms facts from it together with other things. A Russellian proposition is defined as a number of ordered universals and possibly particulars, so that either they combine into a fact or the universal *negation* and they combine into a fact (Hossack 2007: 62ff). Particulars and universals can then be defined in a strikingly Aristotelian fashion. Aristotle writes in the *Categories* that 'a substance is that which is neither said of a substance nor in a substance' (2a13). According to Hossack, 'a particular is anything that occupies "predicate position" in no Russellian proposition; a universal is anything that is not a particular' (Hossack 2007: 66). It follows that facts are particulars.

It is from here that the appearance of abundance in Hossack's account of universals stems. If a universal is anything that is not a particular, and a particular is anything that cannot occupy predicate position in a proposition, then anything that can occupy predicate position in a proposition is a universal, and so it may appear as if any predicate that occurs in a sentence that expresses a proposition refers to a universal. Which propositions there are, however, and hence which sentences express propositions, depends on which facts there are, and which facts there are is determined by nature. If nature is sparse, then so are the universals, and Hossack, as we saw, follows Plato in this respect.

According to Russell, all knowledge must involve universals. It is a further question whether knowledge may also be had of particulars. On Hossack's account, this question is settled decisively. The fundamental epistemological relation is the relation of knowledge between a mind and a fact. Facts belong in the category of particulars, and so do minds. Hence knowledge is a relation between two particulars, one of which is a mind, the other a fact, in that order. Thus, some particulars are known, namely the facts. In *The Philosophy of Logical Atomism*, Russell famously argued for the view that while facts could be asserted to obtain, they could not be named (Russell 1918: 507f): facts are *sui generis* and unlike universals or particulars. Hossack disagrees: as facts are particulars, nothing prevents them from being named.

Knowledge defines the nature of mind. A mind is something that stands in the knowledge relation to some fact at some time. It is of the essence of a mind that it knows something (Hossack 2007: 169). One should maybe add, to allow Socrates or a sceptic to have a mind, that a mind need not know that it knows something. Whether the mental constitutes a special substance, as Descartes held, or is identical to something physical, such as a person's brain, is left open by this account of the mark of the mental. Cartesian souls and brains are equally candidates for the relata of the primitive knowledge relation. However, insofar as knowledge is not a relation known to physics, the physicalist option may only be open to non-reductive physicalists or property dualists.

Central to the philosophy of mind are the notions of consciousness and conscious perception. Hossack explains both in terms of the primitive notion of knowledge. 'Perception is the faculty that gives us knowledge of facts about our environment' (Hossack 2007: 14). Hossack's theory of consciousness builds on an identity thesis which he derives from the Scottish philosopher Thomas Reid: 'A mental act is conscious

if and only if it is identical with knowledge of the quale it itself has' (Hossack 2007: 169). Qualia are often explained as the 'subjective character' of some mental states, in particular of perceptions. Hossack explains qualia also by an identity thesis and a further appeal to the primitive notion of knowledge. 'One's pain is identical with one's awareness of the qualitative character of the pain' (Hossack 2007: 185). A conscious perception is identical with knowledge, or awareness, of the quale of the perception. A quale is a property of an experience. According to Hossack, a quale is a universal Q such that, for any x , if x instantiates Q , then x is identical to knowledge of x (Hossack 2007: 186). Thus conscious perception is a form of knowledge, but the crux of the matter lies in that which is known. Conscious perceptions instantiate qualia, and an instantiation of a quale is identical with knowledge of the instantiation. Suppose Frege smells the scent of violets. This event is the same as Frege's knowing that he smells the scent of violets. His experience has a certain quale, that of smelling of violets, and Frege knows that it does:

e = Frege smells the scent of violets = Frege knows that e has the quale of smelling of violets = Frege knows that Frege's smelling the scent of violets has the quale of smelling of violets.

In general:

X perceives p = X knows that X 's perceiving that p has quale Q .

The perception has the fact that is the perceiving of the percept by the perceiver as a constituent. Thus the fact that X perceives p occurs as a constituent of a constituent of itself, as it instantiates a quale and it is known by X that it does so.

Hossack's theory of consciousness appeals to facts that contain themselves as constituents of some of their constituents. Hossack draws a parallel between this phenomenon with one found in non-standard set theories. In the terminology of set theory, the facts that constitute consciousness are not well-founded. But as non-wellfounded set theory is a respectable mathematical theory, and consistent if standard set theory is, Hossack argues that non-wellfoundedness is not objectionable in the theory of facts that explain consciousness either.

As this aspect of Hossack's philosophy of mind is not as well known as his metaphysics, it may not go amiss to spell out some of the basics of non-wellfounded set theory on which his account of consciousness draws. There is a certain discrepancy in the names given to the axioms of Zermelo-Fraenkel set theory, but with adjustments of terminology, what is to follow may be found in any standard textbook on set theory, such as those of Fraenkel, Bar-Hillel and Levy (1973), Enderton (1977) or Mendelson (2015).⁴

What is commonly referred to as the Axiom of Regularity states that every non-empty set has a member that is disjoint from it: $\forall x (x \neq \emptyset \rightarrow \exists y (y \in x \wedge y \cap x = \emptyset))$. A consequence of the axiom is that no set contains itself. For suppose $A \in A$, and consider its singleton set $\{A\}$, which exists by the Axiom of Pairing. It is non-empty, so by Regularity it has an element that is disjoint from it; i.e. for some y , $y \in \{A\} \wedge y \cap \{A\} = \emptyset$. But the

only element of $\{A\}$ is A , hence $y = A$. But $A \cap \{A\}$ contains A , hence is not $\emptyset$. Contradiction. More generally, suppose there is a finite sequence of sets $A_1 \in A_n \in A_{n-1} \dots \in A_2 \in A_1$. Then by the Axiom of Separation, there is a set $A = \{A_n, A_{n-1} \dots A_2, A_1\}$ containing all and only these sets. A is non-empty, and hence by Regularity it contains an element disjoint from it. But $A_n \cap \{A\}$ contains A_1 , and for any $i < n$, A_i contains A_{i+1} , hence the intersection of A with any of its elements is non-empty, contradicting Regularity. Hence, if facts were like the ordered n -tuples of set theory, Zermelo-Fraenkel set theory would exclude the existence of facts of the kind on which Hossack's theory of consciousness relies. Hossack's facts, therefore, are not such ordered n -tuples.

What is commonly referred to as the Axiom of Foundation states that there is no function that has as its domain the set ω of the natural numbers and as its range a sequence of sets $\sigma_0, \sigma_1 \dots$ such that every i , $\sigma_{i+1} \in \sigma_i$. The Axiom of Foundation implies the Axiom of Regularity, and in the presence of the Axiom of Dependent Choice, the Axiom of Regularity implies the Axiom of Foundation. A consequence of the Axiom of Foundation is that there are no infinite descending sequences of sets standing in the membership relation. Following down the elements of the elements of the elements $\dots$ of a set A , we reach an end after finitely many steps. Any sequence of members of any set A such that $\dots \in x_3 \in x_2 \in x_1 \in A$ ends after a finite number of steps with the empty set. Thus, in Zermelo-Fraenkel set theory, the membership relation is *well-founded*. In non-wellfounded set theory, which rejects the Axioms of Foundation and Regularity, it is not. Infinitely descending and circular sequences of sets standing in the membership relation are permitted. But notice that, just as already finite circles of membership are excluded in Zermelo-Fraenkel set theory, non-wellfoundedness does not always induce infinite chains or infinite circles of membership: if $A \in A$ and nothing else is, the chain is only finite. Some non-wellfounded sets have only finitely many elements and are perfectly finite objects. This important detail forestalls the objection that Hossack's theory of consciousness should be rejected because it induces some kind of 'infinite regress' in the non-wellfounded facts to which it appeals. This is not the case. The facts on which Hossack's theory builds may be as finite as any well-founded facts.

Rationalism is the doctrine that there are truths about the world that can be discovered by pure reason alone, that there is *a priori* knowledge about mind-independent reality. An unwavering belief in the power of pure reason is characteristic of Hossack's approach to philosophy. If pure reason leads us to conclusions that are unorthodox, counterintuitive or uncommon, then so much the worse for orthodoxy, intuition or common opinion. It is also typical of Hossack that he is capable of approaching a discussion afresh, throwing new light on it by looking at it in rationalist fashion from first principles and following a line of thought where it leads.

One example concerns Hossack's views on the relation between necessity and the *a priori*. According to Hossack, the truths that are knowable *a priori* are exactly those that are necessary. This goes contrary to the received view, propounded in Kripke's *Naming and Necessity* (1980), that there are necessary truths which can only be known *a posteriori*, such as that the Morning Star is identical to the Evening Star or that water is H_2O . Hossack's line of argument to establish that the necessary and the *a priori* coincide is elaborate and intricate, and we refer the reader to Hossack's writing on the matter (Hossack 2007, ch. 4), which provides a refreshing challenge to received opinion.

In the philosophy of mathematics, Hossack (2020) shows how one can have an *a priori* science of mathematics. His philosophy of mathematics revives the Aristotelian thesis that numbers are magnitudes, and thus a kind of property. More precisely, a magnitude is a property that is shared by equivalent quantities. For example, the natural number 2 is a property that is shared by all pairs. This simplified version of the Magnitude Thesis plays a substantial role in Hossack's project.

Hossack starts from the Aristotelian thesis that the subjects of a judgment are divided into the categories of *individual* and *quantity*. Examples of quantity considered by Hossack are pluralities (some books), continua (such as a stretch of space and time) or series (such as Plato and Socrates in that order). According to Hossack's metaphysics of mathematics, the natural numbers are properties of pluralities, the positive real numbers are properties of continua, and the ordinal numbers are properties of series. These different kinds of numbers reflect different categories of quantity. Hossack's project thus runs in an anti-Fregean direction. According to Frege and his neo-Fregean followers, numbers are particular objects, which are the referents of semantically singular terms.

Hossack pursues the Aristotelian line of thought that a quantity is what is divisible into 'two or more constituent parts' (Aristotle, *Metaphysics* 1020a7). Thus, quantities have a mereological, or part-whole, structure. Hossack shows that the axioms for the natural and real numbers, and also for the ordinal numbers, can all be deduced from mereology plus the *a priori* axioms that Euclid calls the Common Notions:

1. Quantities which are equal to the same quantity are also equal to one another.
2. If equals are added to equals, the wholes are equal.
3. If equals are subtracted from equals, the remainders are equal.
4. Quantities which 'coincide' with one another are equal to one another.
5. The whole is greater than the part.⁵

Thus, quantities satisfy a notion of equality. For example, two pluralities are equal just in case they can be put in one-one correspondence. In Chapter 4 of his *Knowledge and the Philosophy of Number* (2020), Hossack discusses Leśniewski's mereology and its axiomatization by Tarski. He argues that W. V. Quine, Peter Simons and David Lewis have interpreted mereology as a theory about 'individuals', and thus the axioms of mereology, as interpreted by them, cannot be known *a priori*. According to Hossack's alternative interpretation, mereology is a theory about 'items in the category of quantity' (Hossack 2020: 51).

Hossack puts forward nine axioms of mereology each of which is known *a priori* when interpreted as a law of the logic of pluralities. Likewise, he puts forward nine axioms of mereology each of which is known *a priori* when interpreted as a law of the logic of continua. He shows that eight of the nine axioms are common to pluralities and continua. Hossack proves that these Common Axioms are deductively equivalent to Tarski's and Simons' axiom systems, and concludes that this technical result mandates an interpretation of mereology as the 'pure *a priori* logic of quantity' (2020: 7).

The upshot of Hossack's mathematical project is that our epistemology of number neither relies on empirical evidence nor on set-theoretical authority. It does not appeal

to Fregean abstractionist resources, either. Hossack argues that Euclid's Common Notions and the fundamental laws of mereology are evidently *a priori*, and so, 'by reason alone', we can arrive at our knowledge of number and of number-theoretic truths.

Contributions

The first chapter is a short statement of **Keith Hossack's** philosophical views, which he wrote specifically for this volume, reflecting the main theses and developments of his thought. We will let it speak for itself.

Mark Sainsbury's contribution, 'Confronting Facts: on Hossack's *The Metaphysics of Knowledge*' (Chapter 2), concerns a central theme of Hossack's statement: *relation dualism*. This is the doctrine that there is an irreducible, primitive term in the theory of mind, which refers to something that is both mental and a relation. Hossack holds this doctrine to be true and calls this relation 'awareness' or 'knowledge of'. He takes the notion to be unanalysable and metaphysically fundamental.

Sainsbury does not take issue with the concept of knowledge being unanalysable, but finds it 'harder to say what metaphysical fundamentality amounts to' (p. 39). One natural way of understanding it is in terms of the directness or simplicity of the way the subject relates to the world. In a memorable example of Hossack's, this is the simple and direct way in which an amoeba is aware of the light.

According to Sainsbury, there is a tension in Hossack's views on knowledge. Hossack holds both (a) that knowledge is metaphysically fundamental, and (b) that knowledge is caused by belief under suitable conditions. That knowledge is caused by belief, where belief is a relation to a proposition, would make knowledge appear non-fundamental. For, first, any knowledge-of – the direct, fundamental sort of knowledge – would come causally downstream of knowledge-that – the indirect, non-fundamental sort. And so knowledge-of would be dependent on knowledge-that. Whereas, Sainsbury takes it, that which is metaphysically fundamental is that which is non-dependent. Second, knowledge-that is mediated by a proposition – but how can knowledge-of, which is simple and direct, be caused by knowledge-that, which is a mediated relation to reality?

Sainsbury offers Hossack a friendly amendment: a way of giving knowledge-of a 'special metaphysical and explanatory role', a role he takes Hossack, in any case, to be committed to. Sainsbury does so by working out the view that perceptual experience can be a 'distinctive source of knowledge-of which is not caused by knowledge-that' (p. 40). According to his theory, (i) perceiving is gaining knowledge-of tropes without any specific knowledge-that, and (ii) non-inferential cognitive processes can select 'one of possibly many ways of conceptualizing that input to deliver knowledge-that' (p. 43). This is, roughly, how knowledge-of can yield knowledge-that, in the case of perception.

Sainsbury takes this to explain how knowledge-of – something which is direct, and so non-propositional and obtained non-inferentially – can be a source of knowledge-that, which is propositional, and so can play a role in inference. He takes an important merit of his theory to be its ability to make sense of our experience of perception – as

distinct from testimony – which we express largely through metaphors: open, vivid, rich, coercive, immediate, present, simple and direct.

For Hossack, knowledge consists in a relation between a mind and a fact, and a fact always contains some universal. Hossack's realism about universals is a kind of Platonism. His epistemology resembles a traditional reading of Plato's, on which knowledge has as its subject matter the Forms. In her chapter, 'Who Knows' (Chapter 3), **MM McCabe** argues that there are significant differences between the approaches of Plato and Hossack, and contrasts Hossack's account with an epistemology developed from Plato's dialogues, in particular the *Republic*. In a nutshell, on Hossack's account, knowledge may be had of a very small number of facts and by a mind just happening to stand 'passively' in the right kind of relation to facts, whereas for Plato, knowledge is always of a large range of facts and can only be acquired actively.

On Hossack's view, it should be possible to know only very few facts. If I stand in the right (or wrong, depending on how you look at it) causal relations to the world, I may know that Jamina is a pygmy hippopotamus and not much else. To know this fact I also need to know a few others, such as at which zoo I saw her and that she is a fairly large, round animal with greyish, brownish skin, who is mostly lying in her pond or eating large amounts of hay. But it certainly does not require much knowledge about pygmy hippos, not even that she is a member of a species that occurs in the wild only in West Africa. Furthermore, it is not much of my doing that I came to know that Jamina is a pygmy hippo: it sufficed that we both were in the right place at the right time and causally connected in the right way.

This account of knowledge contrasts sharply with Plato's, explains McCabe. According to Plato, knowledge is systematic, it is a virtue, it is the foundation of wisdom and it is good. Knowledge encompasses entire domains of what is known. Thus, it can only be acquired in a laborious way through long years of training and study. McCabe's Platonist epistemology sees knowledge as akin to understanding, which brings with it many related items of knowledge at once. Understanding is the result of learning, and learning is a process in which an agent engages. Being a virtue, to know requires a certain kind of agent. Just as 'doing the right thing' does not make you virtuous, 'having the right state of mind' does not make you knowledgeable. Being the foundation of wisdom, knowledge is hard to acquire. The value of knowledge lies in the development of the knower. For Plato, to know requires conditions internal to the agent that Hossack's externalism neglects, argues McCabe. It requires an active process in which the knower engages, not the passive one of a mind standing in the right causal relations to facts. The path to knowledge is long, arduous and essential to what it is to know. It also requires an appreciation of one's own epistemic state: to know, in the words of the god at Delphi, you need to know yourself.

Despite their differences, there is an illuminating agreement between Hossack and McCabe. Both agree that knowledge is not a species of belief. Coming to know is moving from *doxa* (belief) to *epistēmē* (knowledge) which are different powers or faculties, as McCabe puts it: they are two different states of mind.

Scott Sturgeon's chapter 'Knowledge-first Epistemology and the Input Problem' (Chapter 4) takes a very different approach from that of McCabe and represents epistemology in the most abstract, theoretical way possible. Sturgeon describes what he

calls the 'Input Problem', from which he takes a number of epistemologies to suffer, including Hossack's knowledge-first epistemology.

Sturgeon outlines what he calls the 'full-dress theory of epistemic rationality', which is composed of four parts: a theory of epistemic targets (which says what is subject to epistemic evaluation), a theory of evidential information (which specifies which pieces of information count as evidence for a situated agent), a theory of evidential support (which explains how a situated agent comes to possess particular bodies of evidence) and a conversion theory (which says how agents should configure their epistemic targets in light of their overall situation).

Sturgeon maps out the logical space open to theories of evidential information and theories of evidential support. In both cases, the theory can be characterized in terms of whether it accepts or rejects each of three components; and in both cases, only six of the eight resulting options are viable. For theories of evidential information, these components correspond to answers 'Yes' or 'No' to the questions whether that from which evidence arises has propositional content, whether it supervenes on the mental, and whether evidence is veridical. For theories of evidential support, the questions are whether an agent's evidential support may be had *a priori*, whether there can be more or less of it, and whether it is absolute or relative to an agent.

Under certain plausible assumptions about the relation between knowledge, evidence and credence, Sturgeon argues that the credence one should afford to anything one knows must be certainty. But this, he points out, is absurd: an agent may know without being certain, and indeed, knowledge may even preclude certainty, due to the fallibility of agents or sources of knowledge. That is to say, we may know things without being sure. Indeed, in many cases we know things and *ought not* to be sure of them, because we know them on the basis of sources we know to be suspect.

This is what Sturgeon calls 'the Input Problem'. It poses a theoretical difficulty for any epistemology which endorses the idea that evidence comes in bits of information, i.e. what Sturgeon calls 'alethic chunks' offered by sources of information. And for Sturgeon, this includes Hossack's knowledge-first epistemology. He goes on to investigate plausible responses to the Input Problem for a knowledge-first epistemology. This will mean rejecting one of the assumptions that generates the trouble: (a) that knowledge yields pieces of evidence, or (b) that pieces of evidence are always maximally supported by total evidence, or (c) that credence for a claim should always match that claim's evidential support on the total evidence.

In his chapter 'Perceiving X = Consciousness of Perceiving X. Hossack and Brentano on the Identity Thesis' (Chapter 5), **Mark Textor** discusses Hossack's theory of consciousness with attention to a thesis of Brentano's and Hossack's (2006) reading thereof. As Textor touches upon aspects of Hossack's philosophy not covered by other chapters in this volume, we'll devote a little more space to its discussion.

Hossack, Reid and Brentano agree on several points: that in conscious perception, knowledge of the perception and the perception necessarily coincide; that this must be recognized and explained by any satisfactory theory of consciousness; and that any theory that allows for them to come apart must be rejected. If I know that I am in pain or am aware of being in pain, I'm in pain. It is the converse that is striking: if I'm in pain, I have immediate and direct knowledge of the pain. The thesis that consciousness of

perception and perception are identical immediately accounts for this remarkable feature. There are not two events that always occur at the same time. There is only one. Conversely, explains Textor, Brentano observes that if the consciousness of a perception and the perception were distinct, the immediate and infallible knowledge we have of our own perceptions would no longer be explicable and, in fact, might fail: if there were two distinct events, each may occur without the other. Thus, Brentano begins his investigation with the epistemology of conscious perception, and its metaphysics must follow the epistemology.

According to Brentano's identity thesis, the mark of conscious perception is that the perception of an object is identical to the perception of the perception of the object. Conscious perception is perception of perception, but perception of perception just is perception. Hossack rejects Brentano's version of the identity thesis, because on his reading of Brentano, he cannot have held that the perception and the perception of the perception are literally the same event. An event, according to Hossack, is a temporal fact. A fact consists in the instantiation of universals by particulars or other universals. An event is, therefore, an instantiation of universals by particulars or universals at a time. Facts are identical if their constituents are identical. According to Hossack's reading of Brentano, a conscious mental event is identical to an event that contains itself as a constituent. But this is impossible, according to Hossack's criterion of identity for events. Consider:

X perceives p = X perceives X's perception of p .

The event on the left-hand side of = contains the perceiver, the percept and perceiving: the perceiver and the percept (in that order) instantiate the universal *perceives* at the time of the perception. The event on the right of = contains the perceiver and the event on the left of =. But the percept p is not identical to X's perception of p . Hence the two events are not identical.

On Hossack's scheme, by contrast, an event that is a conscious mental act does not contain itself as a constituent directly, but only as a constituent of one of its constituents. According to Hossack, a conscious perception is an event e that is identical to X's knowing that e instantiates a quale Q . Then e has as a constituent, not e itself, but the fact that e is Q . It is the latter that has e as a constituent, and although a certain circularity arises, this does not contradict Hossack's criterion of identity of facts (Hossack 2007: 185). The 'circularity' is the kind one encounters in non-wellfounded set theory and, as such, Hossack suggests, it is unproblematic.

To round off the discussion of non-wellfounded set theory, let us go through a detail of non-wellfounded set theory mentioned by Textor. Textor refers to the simplest non-wellfounded set as the one that is identical to its own singleton, i.e. the set A such that $A = \{A\}$. It follows from two basic axioms of set theory, the Axioms of Extensionality and Pairing, and first-order logic that A contains only itself. The Axiom of Extensionality states that if sets A and B contain the same elements, then they are identical. The Axiom of Pairing states that for any set A and B , there is a set containing just them, i.e. the pair set $\{A, B\}$. The case where A and B are identical guarantees the existence of the singleton set $\{A\}$ of any set A , i.e. the set containing A and nothing else:

$$(1) \forall x . x \in \{A\} \leftrightarrow x = A$$

We show that A is the sole element of itself if and only if A is identical to its own singleton. For the left to right direction, suppose A is the sole element of itself:

$$(2) \forall x . x \in A \leftrightarrow x = A$$

So if $y \in A$, then $y = A$, hence y is an element of the singleton of A , i.e. $y \in \{A\}$. Conversely, if $y \in \{A\}$, then $y = A$, hence by (2) $y \in A$. Hence by the Axiom of Extensionality, $A = \{A\}$, and A is identical to its own singleton. Even more briefly, it follows from (1) and (2) by replacement of equivalents that $\forall x . x \in A \leftrightarrow x \in \{A\}$, and so by Extensionality $A = \{A\}$. For the right to left direction, suppose A is identical to its own singleton, i.e. $A = \{A\}$. Then by Leibniz's Law

$$(3) \forall x . x \in A \leftrightarrow x \in \{A\}$$

(2) follows from (1) and (3) by replacement of equivalents, and A is the sole element of itself. Having appealed only to Extensionality and Pairing, $(\forall x . x \in A \leftrightarrow x = A) \leftrightarrow A = \{A\}$ holds in Zermelo-Fraenkel Set Theory, but both sides of the principal biconditional are false, as the Axiom of Foundation precludes the existence of a set that is its own sole member and that is identical to its own singleton.

Textor defends Brentano against Hossack's criticism, and argues that their views are closer than might appear. Textor recommends, however, not identifying mental acts, as understood by Brentano, with Hossack's ordered universals, particulars and times. It is worth noting that Hossack adopts his account of events because 'it fits conveniently with [his] metaphysics of facts' (Hossack 2007: 102): events are just another kind of fact. Another popular account, namely Davidson's, according to which an event is a particular of a more mundane kind, may be better suited to Brentano's metaphysics of mental acts. Further considerations may convince Hossack to adopt it instead.

Textor's defence of Brentano appeals to his most famous thesis: intentionality is the mark of the mental. Mental acts are distinguished from everything else by being directed upon objects. Textor combines it with Hossack's work on plural reference to shed light on Brentano's account of consciousness. Plural reference occurs if an expression refers to more than one object, to a plurality considered as several, not as a unified collection. When I say 'The geese are noisy', I'm not saying that the pair set containing the two Egyptian geese outside my window are noisy. That is an abstract object and accordingly very quiet. It is the geese who are noisy, and 'they' refers to both of them together, as a plurality.

Textor explains that Brentano's mark of the mental should not be understood as if each mental act is directed upon only one object. Mental acts can be directed upon pluralities, just as words can refer to them. In fact, conscious mental acts are always directed upon more than one object, as they are always directed upon the acts themselves. A conscious mental act, such as smelling the scent of violets, is directed towards itself as well as towards some other object, in this case the violets. According to Textor's defence of Brentano, although in conscious perception, the perception and

the perception of the perception are identical, there are two ways of conceptualizing that one mental event, once as a perception and once as a perception of a perception. Thus, the two sentences:

1. Frege smells the scent of violets.
2. Frege is conscious of the smell of violets.

constitute two different descriptions of the same event. The plurality of the objects towards which the mental act is directed – the objects of the perception as well as the perception itself – underlies the different ways of describing it.

Hossack's theory of knowledge, as we have noted above, takes knowledge to be an unanalysable and metaphysically fundamental relation between minds and facts. And one of the many phenomena that can be understood in terms of knowledge is modality. **Bernhard Weiss's** chapter 'Facts, Knowledge and Knowledge of Facts' (Chapter 6) puts pressure on each part of Hossack's theory. First on its ontology of facts; second, on the evidence for treating knowledge as a relation to a fact, and third, on the use of knowledge to explain modality.

First, ontology. What is known, in Hossack's theory, are facts. While facts are metaphysically basic, we can comprehend them through a theory of vectors and combination. Vectors are any entities taken in a particular order. Hossack defines a notion of sense for vectors: the entities of a vector make sense if and only if there is either a fact combining them or there is a negative fact combining them. So combination is a crucial relation in Hossack's understanding of facts. Weiss argues that this ontology of facts requires a *singular* combinatorial relation, whereas a closer look reveals a plurality of relations which relate entities in a vector such that they make sense. Weiss considers the facts that Cassio loves Desdemona and that Desdemona does not love Cassio. According to Hossack, the first fact combines the universal *love*, Cassio and Desdemona, in that order; the second fact combines the universal *negation, love*, Desdemona and Cassio, in that order. The combination relation treats all items of the vector except the first one as being on a par. The first item relates all the others in the given order. Thus, in the first fact, *love* relates Cassio and Desdemona; and in the second *negation* relates *love*, Desdemona and Cassio. But that, Weiss argues, cannot be correct. Even though there is no fact in which *love* relates Desdemona and Cassio, *love* remains a relation even in the negative fact that Desdemona does not love Cassio. But Hossack must deny that there is anything relational about *love* in this fact. The best we might do, says Weiss, is to accept that negation is a further combination relation, rather than a universal, responsible for those facts that are negative. But this solves the problem at best for the case of *negation*, not for any other cases that might arise. And the problem is, Weiss argues, that there is an indefinite number of such other cases. In fact, the problem is one Russell had already faced with respect to propositions containing two verbs, amongst which are propositional attitude reports. There shouldn't be a special combination relation for each of those cases. Weiss concludes that due to the proliferation of combination relations, there is reason to doubt that we have a grasp of combination, and this threatens our understanding of what facts are supposed to be in Hossack's ontology.

Weiss's second argument concerns the evidence for knowledge being a relation between mind and fact. Even if one supports Hossack's anti-sceptical *motivation* which leads him to cut out epistemic intermediaries between the knower and that which is known, Weiss argues that the evidence Hossack produces for his way of doing so is weak. It is the factivity of knowledge that does the work here. True propositional attitudes are not factive, and so they are relations to contents, not facts. Knowledge works differently. And so the logic of belief attribution is very different from the logic of knowledge attribution, and the reason for this is that belief and knowledge are very different relations, and factivity can be explained by the referentially transparent relation of mind to fact. Hossack's evidence for the claim that knowledge is a relation between mind and fact relies on the following linguistic argument. Knowledge-that claims entail knowledge-of claims: 'S knows that Fred is red' entails 'S knows of Fred's being red'. Sentences of the latter form require a sentential nominalization for their completion, and they are referentially transparent. S can only know of Fred's being red if Fred is red, and so here we have a direct relation between a mind and a fact. But there is no corresponding relation between a believes-that claim and a believes-of claim. Indeed, there are no believes-of claims at all.

Weiss isn't taken with this argument and observes that similar entailments hold for the propositional attitudes 'believes', 'hopes' and 'fears'. While these involve slightly different grammatical constructions, Weiss questions whether they are of logical significance. But as in these examples, the propositional attitudes must report a relation between a mind and a content, and so Hossack's linguistic argument fails to support his conclusion that knowledge is a relation between a mind and a fact.

Weiss's third argument is that Hossack's attempt to explain modality in terms of knowledge is implausible – a point on which he and Dorothy Edgington (Chapter 8) agree despite critiquing different versions of Hossack's argument for the same conclusion, Weiss the old version (Hossack 2007), and Edgington the new version published in the present volume.

According to Hossack's rationalist thesis, necessary facts coincide with those that are knowable *a priori*, that is, with those that have an *a priori* mode of presentation. Thus both notions, necessity and apriority, apply to facts. But, Weiss points out, apriority is usually taken to apply to contents, not facts. Hossack's fellow rationalists Leibniz, Kant and Frege understood both notions in this way: for them, necessity and apriority apply to contents. Their coincidence faces the usual counterexamples. There are sentences that are known *a priori* but that are not necessarily true, such as those arising when a name is introduced by stipulation – e.g. Kripke's famous example of 'Neptune' (Kripke 1980: 79, fn. 33) – and there are sentences that are necessarily true, but that are known *a posteriori*, such as sentences expressing identities – e.g. 'Hesperus is Phosphorus' (Kripke 1980: 104). Hossack's account may not succumb to those counterexamples, but Weiss discerns a change of subject. Whereas Hossack gives an account of *a priori* knowledge-of, we should expect an account of *a priori* knowledge-that. Weiss concludes that there is good reason to accept that apriority is a property of knowledge-that, and knowledge-that cannot explain necessity. Hence Hossack's account fails.

Hossack doesn't merely treat knowledge as unanalysable and metaphysically fundamental – in the sense problematized by Sainsbury. Part of putting 'knowledge

first' means explaining other things in terms of it, that is, exhibiting its explanatory fruitfulness (Hossack 2007: xi–xvi). And, as we have just seen, part of the explanatory power of knowledge, for Hossack, is that it can be put to work in a reductive account of modality. This can be seen as the core tenet of his rationalism: Hossack's *rationalist thesis* is that 'a fact is necessary if and only if it has an *a priori* mode of presentation' (Hossack 2007: 125).

In 'Necessity, Conditionals and Apriority' (Chapter 7), **Keith Hossack** aims to prove that the propositions we can know by reason alone are necessary, and that the necessary propositions are those that are knowable *a priori* – in short, that the necessary and the *a priori* coincide. Given that necessity and possibility are interdefinable, Hossack takes himself to have 'reduce[d] the modal to the epistemic' (p. 125).

Part of the interest of this position is that it violates a dogma enforced by followers of Kripke: that the *a priori* and the necessary do not coincide. This idea renewed interest in metaphysics, for it gave philosophy a new task: to reveal the broadly *a posteriori* necessities. This tied metaphysics closely to science in two ways. The results of science could amount to significant metaphysical (not merely physical) revelations, such as the identity of water and H₂O. And so too the methods of science could be harnessed for metaphysics, as Lewis infamously used inference to the best explanation to 'discover' the pluriverse (Lewis 1986).

Hossack hopes, by contrast, to vindicate the traditional pre-Kripkean idea that the *a priori* and necessary coincide – but without going down the positivist route and identifying the necessary with the analytic. He takes the coincidence of necessity and apriority as falling out of Nelson Goodman's definition of a counterfactual statement, according to which a counterfactual is true if its consequent is inferable from its antecedent (Goodman 1979).

Hossack's argumentative strategy is as follows:

1. Prove Goodman's truth-conditions for counterfactuals.
2. Establish the definition of necessity in terms of inferability.
3. Establish the definition of apriority in terms of inferability.
4. Prove that Goodman's truth-conditions for the counterfactual entail that necessity and apriority coincide.

This approach stands in stark contrast with David Lewis's, the best known, if not the most widely believed, reduction of modality. Lewis wanted to reduce the totality of modal discourse to quantifications over the pluriverse, where the pluriverse is all the concrete possible worlds (Lewis 1986: 7); so necessities become truths about what goes on in *all* such worlds, and possibilities become truths about what goes on in *some* of them. In Lewis's picture, there is no essential connection between apriority and truths about the pluriverse. Hossack, on the other hand, through the employment of the conceptually unanalysable and metaphysically fundamental notion of knowledge, sees an essential connection between mind and world, such that the necessities are available to knowledge.

Hossack's paper is put under the knife by **Dorothy Edgington** in her commentary on it (Chapter 8). She presents a cumulative argument whose conclusion is that 'I don't

feel forced to the conclusion that the necessary and the *a priori* coincide' (p. 141). Nor, by implication, does she think we should be either. She takes issue with the central notions of Hossack's attempted reduction of the modal to the epistemic.

Edgington takes the core notion of Hossack's account to be obscure. This is his notion of the *far-fetchedness* of a proposition which he employs instead of that of *closeness to actuality*. He uses *far-fetchedness* to 'complete' Goodman's definition of the counterfactual, such that it is non-circular and does not rely on an undefined or Lewisian notion of a possible world. *Far-fetchedness* is itself defined in terms of a counterfactual, thus making little advance on Goodman's own circular definition of *co-tenability*. Likewise, Edgington finds the notion of propositional *relevance* to be problematic.

Further, Edgington finds Hossack's proof of his version of Goodman's theory of counterfactuals to be convincing only for those who already accept that 'an ordering by closeness/far-fetchedness is a key to the interpretation of counterfactuals' (p. 138). But she shows that such an ordering is not 'primitively evident' and that there is in fact substantial disagreement by competent philosophers on whether an ordering is 'a key to the interpretation of counterfactuals' and what the basis of the correct ordering might be.

Edgington is unconvinced by Hossack's final proof of the coincidence of necessity and apriority. According to Hossack, if *P* is necessary, then there is a proof of *P*. But as Edgington points out, as a consequence of Gödel's first incompleteness theorem, not every true proposition of number theory has a proof in the formal system of Peano Arithmetic. For instance, Goldbach's Conjecture is unproven. If it is true, it is necessarily true. It is counterfactually robust: whatever else were the case, it would be true. Thus, for Hossack, as it is *a priori*, it should be provable from self-evident premises. Many will agree with Hossack that an adequate notion of proof outstrips what is provable in the formal system of Peano Arithmetic: there are more ways of establishing propositions as true than there are ways of proving formulas in Peano Arithmetic. This is often taken to be a philosophical consequence of Gödel's Second Incompleteness Theorem and also of Tarski's Theorem about the undefinability of the truth predicate for arithmetic within Peano Arithmetic. But Hossack accepts that some proofs are infinitely large, such as proofs that appeal to the ω -rule. This rule requires infinitely many premises: it permits the derivation of the conclusion $\forall xFx$ from the infinitely many premises Fn , for every n . Goldbach's Conjecture may well have a proof in this sense (if it is true, calculate each instance and apply the ω -rule). But, Edgington points out, this goes against what most would ordinarily require of a proof, namely, that it enables us to come to know its conclusion on the basis of our knowledge of the premises. This is not possible in a proof that appeals to the ω -rule, as we cannot come to know its infinitely many premises.

Edgington also provides counter-examples to the other direction of Hossack's equation of the necessary with the *a priori*, suggesting that there are *a priori* truths which are not necessary. For instance, 'What Dee weighs at t , is her actual weight at t '. So neither direction of the supposed coincidence is, for Edgington, plausible.

The thrust of Edgington's observations, counterexamples and arguments is that Hossack has not established what he had hoped: that the necessary and *a priori* coincide. And so the modal has not yet been reduced to the epistemic.

Questions relating to *a priori* knowledge and mathematical proof are further pursued by Tamsin de Waal in her chapter 'The Mathematicians' Use of Diagrams in Plato' (Chapter 9). *A posteriori* knowledge is gained through experience. But how do we gain *a priori* knowledge? Mathematics is often cited as providing the paradigm case of an area of inquiry where *a priori* knowledge is to be found. Knowledge of mathematical propositions is usually acquired through proof. In geometry, these proofs often do, and in ancient Greek times essentially did, involve diagrams. But a diagram depicts a particular situation, while mathematical propositions are general. De Waal's central question is: how do we apprehend general mathematical propositions through diagrams in the proofs of geometry? How can a diagram that shows a particular triangle enable us to apprehend a general truth about all triangles? Geometry, as de Waal reminds us, was paradigmatic of all of mathematics in ancient times, and so her question was particularly pressing for Plato and his contemporaries. De Waal's chapter concerns Plato's conception of mathematical knowledge, but her discussion draws general consequences for the epistemology of mathematics, especially the cognitive significance of images and diagrams in gaining knowledge of mathematical propositions.

De Waal begins with a puzzle from Plato's *Republic*: while images are denigrated by Socrates and his friends, the way mathematics was done in Plato's time relied heavily on the use of images in diagrammatic proofs, and the study of mathematics plays a key role in the education of the philosopher-rulers. The denigration of images is a theme at various points of the *Republic*. In the image (!) of the Divided Line – Plato's representation of a hierarchy of four epistemic states at the end of book VI of the *Republic* (510dff) – images are relegated to the lowest epistemic state, that of *eikasia* or imagination. A common resolution of the puzzle is to take Plato as holding his contemporaries to account for using images in mathematical proofs, and finding mathematical practice defective where it does so. De Waal proposes an original alternative solution by arguing for the opposite: Plato endorses the use of images, where these are diagrams, as a fundamental and necessary feature of mathematics. They play an essential part in attaining the higher epistemic states of the Divided Line by facilitating the conversion from *pistis* to *dianoia* – translated by de Waal as *belief* and *understanding* – which is itself an essential step to reaching the goal of *noësis* (rendered by de Waal as *intelligence/dialectic*).

De Waal's solution thus permits a full endorsement of Plato's comments that the study of mathematics *as it is* leads to 'a conversion and turning about of the soul' (*Republic* 521c) and will 'draw the soul away from the world of becoming to the world of being' (*Republic* 521d). This is possible despite the dubious role played by images in the *Republic*, because, argues de Waal, the diagrams of geometrical proofs prompt thought to direct itself towards the intelligible entities and force the recognition of the fact that physical/visible objects are distinct from the intelligible entities and fall short of these in various ways. The positive role played by diagrams beyond the fact that they are a defining feature of the mathematical method lies in the way that the particular situation in a diagram represents a general truth of mathematics, and thus they point to the forms because they lead away from the particular to the general.

A feature that singles out mathematical propositions is their generality. This theme is picked up by **Nils Kürbis** in his chapter ‘Generality’ (Chapter 10). Kürbis offers Hossack three arguments for his thesis that general facts are general in virtue of containing a special universal. To describe the world correctly, we need to use general statements. An important class of them contains those expressing the natural laws. As is the case with any other kind of statement, some general statements are true, some are false. In virtue of what are such statements true or false? Hossack rejects the account put forward by Wittgenstein in the *Tractatus*, according to which general statements are made true by their instances (Hossack 2007: 68ff). This account was also the one held by Russell and Whitehead in the first edition of *Principia Mathematica* (1910: 47). Hossack concludes that, instead, general statements are made true by general facts. These are facts that contain the universal *generality* in principal position. *Generality* is the referent of the universal quantifier of formal logic. Hossack does not, however, provide a sustained and detailed argument for the existence of general facts comparable to his argument for the existence of negative facts, those containing the universal *negation*. Kürbis’s three arguments aim to fill this gap.

All three arguments are inspired by Russell’s writings. The first two draw on Russell’s analysis of the nature of universals as aspects of resemblance. All universally instantiated universals resemble each other in this respect. Hence, we should expect that there is a universal which is that resemblance. This is the universal *generality*. All existential facts resemble each other in this respect, so here, too, we should expect that there is a corresponding universal, namely *existence*. Together the universals *existence* and *negation* can account for generality; conversely, together the universals *generality* and *negation* can account for existence. Given the universal *negation*, we need not assume that *existence* and *generality* are both primitive, and Hossack decides in favour of *generality*.

The third argument follows a famous argument of Russell’s that there must be universal facts, because any inference from purely particular premises to a general conclusion is invalid. Russell, however, was not explicit about the nature of general facts and cautious about what the right analysis of general facts might be. Kürbis argues that although Russell’s argument leaves the nature of general facts open, the previous two arguments settle this question in Hossack’s favour: we should draw Hossack’s bold conclusion that what it is that makes a fact general is that it contains the universal *generality* in principal position.

Simon Hewitt’s chapter, ‘We Belong Together? A Plea for Modesty in Modal Plural Logic’ (Chapter 11) takes issue with one of Hossack’s most important contributions in metaphysics and the philosophy of logic: atomism and plural reference. In his influential paper ‘Plurals and Complexes’ (2000), through the conceptual resources of plural logic, Hossack provides a reduction of our talk of complex objects to the pluralities of their parts. In effect, he argues for atomism, the view that there are no complex things, only simples, the atoms, and our ways of referring to the simples plurally.

Hewitt challenges an important thesis known as Plural Rigidity, according to which if a given thing is one of some given things, then it is necessarily so, given the existence

of the relevant things. If Plural Rigidity is accepted, then Hossack's project of reducing our talk of complex objects to talk of pluralities of their parts seems to be undermined: by Plural Rigidity, complex objects will be necessarily identified with their actual parts. Thus, we will have a version of 'mereological essentialism'; the thesis that for any complex object x , if x has y as one of its parts, then y is a part of x in every possible world in which x exists. But mereological essentialism is often taken to be an implausible claim, for it suggests that to be part of something is to be part of its essence. It seems implausible to claim that, given that my bicycle has a particular wheel as its part, then in every possible world in which my bicycle exists it has that particular wheel as its part.

The Plural Rigidity thesis has received considerable attention in the recent literature. (See Linnebo (2016), Rumfitt (2005), Uzquiano (2011) and Williamson (2013).) The view has been challenged in Hewitt (2012). The rigidity of plurals has played an important role in a number of debates in metaphysics, philosophy of logic, philosophy of mathematics and of language – for example, in the interpretation of higher-order logics; in Williamson's arguments for *necessitism*, the thesis that necessarily everything necessarily exists; and in Linnebo's (2010) programme for interpreting standard set-theory in a modal plural logic.

In his 'Plurals and Modals' (2016), Linnebo defends Plural Rigidity. He argues that Plural Rigidity lies at the core of a number of correct and plausible theses concerning the extensionality of pluralities. Linnebo starts from an analogue of the Necessity of Identity, defended by Kripke (1980),⁶ for pluralities, to the effect that if every one of these objects is one of those objects and every one of those is one of these, then necessarily these just are those. In short: coextensive pluralities are necessarily coextensive:

$$(\text{Cov}) \quad xx \equiv yy \rightarrow \Box (xx \equiv yy)$$

where ' $xx \equiv yy$ ' abbreviates ' $\forall u(u < xx \leftrightarrow u < yy)$ ', stating that whatever is part of the plurality xx is part of the plurality yy , and *vice versa*. ('<' stands for the parthood relation.) As Linnebo puts it, (Cov) tells us that 'two overlapping pluralities necessarily covary in their membership' (2016: 662). But it is not difficult to see that (Cov) does not entail the desired Plural Rigidity, which states that a plurality has the same members at any world in which it exists. Linnebo's strategy for deriving Plural Rigidity is through some further assumptions, to bridge the gap between the extensionality of pluralities and their rigidity.

Hewitt's contribution takes issue with Linnebo's formal arguments in favour of Plural Rigidity. In Hewitt's view, all of Linnebo's arguments suffer from the same type of problem: they fail to persuade somebody not already committed to Plural Rigidity. In fact, the source of the difficulty, in Hewitt's view, lies in Linnebo's starting point: the extensionality of pluralities as codified by (Cov). It is exactly this extensional conception of pluralities that is question-begging.

One of the auxiliary assumptions that Linnebo employs to derive Plural Rigidity is Uniform Adjudication, which states that, necessarily, to be one of these things and that thing is to be one of these things or to be identical with that thing. The proof of Plural

Rigidity rests on (Cov) and Uniform Adjudication. But Hewitt argues that the use of (Cov) in Uniform Adjudication begs the question against someone not antecedently committed to Plural Rigidity. Hewitt makes a similar point concerning the other two arguments discussed by Linnebo.

Hewitt points out that there is also linguistic evidence against Plural Rigidity, but he concedes that such evidence cannot provide us with an effective tool against the use of the principle in our theorizing about the plural concepts we express in our natural languages. In Hewitt's view, in the absence of a philosophical motivation for Plural Rigidity, the appeal to it in support of strong metaphysical theses is illegitimate.

The contributions of **Øystein Linnebo**, **Bahram Assadian**, and **Peter Simons** deal with questions in the philosophy of mathematics, which have direct interactions with Hossack's concerns. As discussed above, according to Hossack, numbers are properties. The natural numbers are properties of pluralities, the positive real numbers are properties of continua, and the ordinal numbers are properties of series. Central to Hossack's project is the Magnitude Thesis, which, to put it roughly, is the thesis that a magnitude is a property that is shared by equivalent quantities. In his contribution, 'Aristotelian Aspirations, Fregean Fears: Hossack on Numbers as Magnitudes' (Chapter 12), Linnebo draws our attention to a structural similarity between the Magnitude Thesis and Fregean abstraction principles. He then explores the question of how different Hossack's magnitudes really are from mathematical objects, as conceived of by Frege and the neo-Fregeans.

Let us say that two pluralities are equivalent just in case the equivalence relation of one-one correspondence holds between them. Thus construed, the Magnitude Thesis involves a function that maps a quantity to its magnitude. Fregean abstraction principles behave in a similar way. Consider Hume's Principle, which states that the number of the concept F (i.e. the number of things falling under the concept F) is the same as the number of the concept G just in case there is a one-one correspondence between the F s and the G s. The cardinality function, denoted by 'the number of', maps a concept to its cardinal number. So, we can put forward a Frege-style abstraction principle for magnitudes as follows:

$$(AP) \quad \phi(x) = \phi(y) \leftrightarrow x \sim y$$

where ' $\phi(x)$ ' stands for the function *the magnitude of*. Thus, (AP) states that the magnitude of the quantity x is identical to the magnitude of the quantity y just in case x and y are equivalent, for some suitably chosen equivalence relation, depending on the sort of quantities x and y are.

In Hossack's view, there is an important reason for favouring his Aristotelian approach concerning the nature of numbers over the Fregean one. The entities on the left-hand side of (AP), the magnitudes of certain quantities, belong to the category of properties, not Fregean objects. And since magnitudes are sorts of entities that could be acceptable for nominalists who deny abstract *objects*, the Aristotelian approach is preferable.

Linnebo, however, points out that Hossack's axiomatization of the mereological structure of quantities, as developed in Hossack (2020, ch. 4), has been formulated in a

type-free language, in which a single type of variable is used to range over individuals, pluralities, continua, and series. And this, Linnebo claims, 'gives us all that Frege and his followers ever wanted when they defended the idea of numbers as objects' (p. 197). All that Frege's famous bootstrapping argument for the existence of infinitely many natural numbers needed is that numbers figure as members of the pluralities whose numbers we are interested in. This condition, however, does not require a substantial metaphysical notion of objecthood. Thus, if Hossack's conception of numbers is distinctively non-Fregean, his numbers must fail to be objects in some metaphysical sense, but it is not clear whether Frege and his neo-Fregean followers need such a metaphysically loaded conception of the objectual nature of numbers.

Linnebo further argues that the abstractionist reading of the Magnitude Thesis in terms of (AP) exposes (AP) to a version of what is known in the literature as the *bad company problem*. According to this objection, abstraction principles cannot be acceptable ways to introduce mathematical concepts, because there is no satisfactory way to separate 'good' abstraction principles, such as Hume's Principle, from 'bad' ones, such as Frege's ill-fated Basic Law V, with which they share their general form. What is the principled account that separates the good from the bad abstraction principles? And how can Hossack respond to the bad company problem?

This question is pressing in view of Hossack's commitment to Unrestricted Comprehension for Series; i.e. the principle that if a two-place formula defines a well-ordering without repetitions, then the formula defines a series. The threat of the Burali-Forti Paradox is familiar here: by Unrestricted Comprehension for Series, there would be a series of all the ordinals, and by Hossack's theory, this series would define an ordinal which would be greater than all the ordinals – including itself. This paradox forms a special instance of the bad company problem.

Hossack's proposal (2020, ch. 10) for blocking the Paradox is to restrict the range of series only to those that are defined by *recursive* well-orderings. Linnebo, however, argues that Hossack's proposal is 'both unacceptably restrictive and ultimately *ad hoc*' (p. 199). It is restrictive because, for example, the uncountable plurality of the real numbers can be well-ordered, but the condition of recursivity prevents us from making serial reference to the real numbers in that order. And it is *ad hoc*, because there is no connection between Hossack's responses to the Burali-Forti Paradox and Russell's Paradox: in the latter case, Hossack makes a Quinean move by appealing to NFU; i.e. a version of Quine's New Foundations with entities which are not sets, known as *Urelemente*. But in the former, Hossack restricts the range of series to recursive well-orderings. The different explanations for similar set-theoretic paradoxes need a strong justification, which, in Linnebo's view, is lacking in Hossack's proposal.

The main focus of Bahram Assadian's chapter, 'Mathematical Structures, Universals and Singular Terms' (Chapter 13) is structuralism in the philosophy of mathematics; the thesis that foundational mathematical theories are about *abstract structures*. But what is an abstract structure? This question is related to Hossack's detailed account of the nature of ordinal numbers (2020, ch. 10), according to which an ordinal number is a universal: it is what isomorphic serial relations, i.e. isomorphic well-ordered series, 'have in common', where two serial relations, R and S , are isomorphic if and only if there is an order-preserving bijection f from the field of one onto that of the other; i.e.

the function f is such that: $\forall x \forall y (Rxy \leftrightarrow Sfxfy)$. An ordinal, thus construed, is a particular instance of an abstract structure, which is what isomorphic relations, and not necessarily well-ordering relations, instantiate or have in common. Assadian's chapter can thus be seen as a way of extending Hossack's conception of ordinals to an account of structuralism in which abstract structures are identified with universals.

In the literature on structuralism, abstract structures are what isomorphic 'models' or 'systems' of a mathematical theory have in common, where a system, following Shapiro (1997), is defined as an ordered $(n + 1)$ -tuple consisting of a domain D and relations $R_1, \dots, R_n$ on this domain. A much-discussed account of abstract structures, known as *ante-rem structuralism*, is defended and developed by Shapiro (1997, 2008). According to this view, an abstract structure corresponding to a mathematical theory is a *structural universal* shared by the isomorphic systems of theory, such that the universal consists of a category of purely structural entities, known as *positions*. The question Assadian deals with is this: could an *ante-rem* structure be the kind of entity the structuralist has been looking for as a candidate for *abstract* structures, namely as a candidate for what isomorphic systems of a theory have in common? Assadian discusses two arguments leading to a negative answer and offers what the structuralist needs in order to resist these arguments.

The first argument is Geoffrey Hellman's (2001: 195–196) permutation objection. Hellman argues that ante-rem structures themselves, just like their instantiating systems, are subject to the sort of indeterminacy Benacerraf (1965) has argued for. Thus, they leave us with distinct and yet isomorphic copies. The reference to positions of structures will be as indeterminate as reference to the elements of their corresponding systems. Referential indeterminacy does not go away despite postulating ante-rem structures.

Assadian argues that the source of the problem of referential indeterminacy lies in the structuralist's conception of positions as *bona fide* objects, as the referents of semantically singular terms. He thus puts forward a thesis, recently defended by Hale and Wright (2012), Hale and Linnebo (2020) and Hossack (2020, ch. 1 and ch. 4), according to which singular terms can refer to universals. Thus, contrary to what Frege recommended, it is not the case that *only* objects can be the referents of singular terms. This will liberate the structuralist from positing a special *system* with a domain of positions-as-objects to serve as the unique referents of our mathematical expressions: we can successfully refer to the universal *natural number structure* by using the singular term 'the structure of natural numbers', as opposed to the systems having this universal in common. The universal *natural number structure* is a structural universal, and we can permute its sub-universals, i.e. its positions, in a way that produces an isomorphic system of that structure. This isomorphic system itself is not the universal *natural number structure*. All the same, we can still use the singular term 'the structure of natural numbers' to refer to the universal rather than to a system obtained from it by permuting its sub-universals. The idea that the singular terms formed by an abstraction principle can be understood as 'derived' reference to universals comes up also in Linnebo's chapter in this volume, where he provides a reconciliation between Hossack's Aristotelianism on the nature of numbers as universals and Frege's abstractionism.

The second argument against the ante-rem conception of abstract structures is due to John Burgess (1999). According to this argument, an ante-rem structure is supposed to be a system, which is an ordered set, and so it is supposed to consist of a collection of positions-as-objects, with distinguished relations on them. But abstract structures as what isomorphic systems have in common – what Burgess calls ‘isomorphism types’ – are abstracts with respect to the equivalence relation of isomorphism. Since the *criterion of identity* for abstract structures and that of ante-rem structures are distinct, the two categories cannot overlap. Assadian argues that Burgess’s argument rests on the conception of ante-rem structures as special systems whose positions are taken to be objects. Again, once we dispense with this objectual perspective, there will be room for the identification of ante-rem structures, taken as structural universals, with abstract structures.

Peter Simons’ chapter, ‘Arithmetic in a Finite World’ (Chapter 14), defends a nominalist account of numbers, and thus sharply contrasts with Hossack’s realism in the philosophy of mathematics. In fact, Simons’ starting point accords well with Hossack’s thesis that numbers are properties – for example, the natural numbers are properties of collections. But Simons defends a radically contrasting thesis concerning these properties. In his version of nominalism, these properties are *fictions*. In particular, Simons offers a nominalistic reading of Hume’s Principle. In his view, numerical expressions of the form ‘The number of *F*s’ are not genuinely referential expressions.

Starting with the natural numbers, Hossack’s thesis faces the familiar problem for accounts which take mathematical statements to be grounded in individuals, pluralities or collections of them: how can it be ensured that there are enough individuals, pluralities or collections to validate the axioms of Peano Arithmetic?

Simons aims to ground arithmetical truths, which require infinitely many objects, in a finite (concrete) world. Frege’s proof of the infinity of the natural numbers, also discussed in Linnebo’s contribution, indispensably depends upon the objectual nature of numbers – at least upon numbers being objects in a general logical sense. But in a Russell-style type theory, the Fregean proof is not available. For Russell, a number is a class of classes, and thus belongs to at least level 3 in the hierarchy, two levels up from level-0 individuals. As a result, Russell has to appeal to the Axiom of Infinity, which asserts that there are infinitely many individuals. However, as Russell was aware, and Simons reminds us again, the logical character of the Axiom of Infinity is highly dubious. Russell writes:

Nor will it do to add the *Axiom of Infinity* as a hypothesis to theorems in whose proof it is used – a procedure which is adopted in *Principia Mathematica*. To have recourse to this tactic is to abandon the logicist project – supposing we had a satisfactory axiomatization of a physical theory, no one would be taken seriously who claimed to reduce that branch of physics to logic by rewriting each theorem of the theory as a conditional with the conjunction of the axioms as antecedent. When the number theorist asserts that there are infinitely many primes, he is making a categorical assertion, and not merely claiming that if there are infinitely many individuals, then there are infinitely many primes.

Russell, 1919c: 141; see also 202–203

Simons reminds us that the main reason for Whitehead and Russell, in their *Principia Mathematica*, to postulate the non-logical Axiom of Infinity was to guarantee the existence of an infinite number of objects. Non-logicality aside, Simons also holds that the Axiom of Infinity has the ‘disturbing consequence that Peano arithmetic may be empirically false, if the world is finite in individuals’ (p. 222).

Given the assumption of the existence of at least two individuals, and also the existence of higher-order collections (pluralities) – collections of collections, collections of collections of collections, and so on – Simons shows, in a type-style hierarchy, how we can capture the axioms of the Dedekind-Peano Arithmetic; hence the title of his chapter, ‘Arithmetic in a Finite World’. Collections of individuals are different from the individuals of the collections. The collection of the Moon and the Sun is not identical to the Moon and the Sun. A collection is not an individual, for otherwise there would be infinitely many individuals. But the mere existence of the Moon and the Sun ensures the existence of their collection. So, we have at least three items if we have the Moon and the Sun: the Moon, the Sun and their collection. But we will also have a higher-order collection: the collection of these three items. Each level, based on two or any finite number of individuals, is finite, but since there is no upper limit to the orders, infinitely many collections appear in the hierarchy.

An interesting metaphysical consequence of Simons’ account is that ‘Peano arithmetic is not necessarily true’ (p. 222), for its truth requires as an additional premise that there are at least two individuals.

Notes

- 1 During the final and vital stages of this project, Nils Kürbis was supported by the Alexander von Humboldt Foundation, to whom many thanks are due.
- 2 One might be tempted to say ‘whenever there are things that are different, there is an aspect in which they do not resemble’. But this requires a commitment to the claim that if *a* and *b* are different, then there is some property that they do not share. In conversation, Hossack has expressed reservations about this half of Leibniz’s Law and pointed out that the desired effect may be had by the weaker claim relying only on dissimilarity.
- 3 In this and the previous paragraph the authors have relied on Hossack’s explanation of his views in private correspondence.
- 4 Nils Kürbis has profited from discussions with Neil Barton in getting to grips with non-wellfounded set theory and would like to thank him here.
- 5 Quoted from Hossack (2020: 4).
- 6 The first published proof that, under certain assumptions, identity statements are necessary is by Ruth Barcan Marcus (1947). The simple proof that this follows by replacement of the Law of Identity in the schematic first-order version of Leibniz’s Law was apparently first published by Quine (1953). See Burgess (2014). The simple proof is also found in Prior (1955: 205f). Prior notes that the simple proof follows from a more complex one, essentially due to Barcan Marcus, if the Law of Identity and the first-order version of Leibniz’s Law are taken as axioms rather than derived from Leibniz’s second-order definition of identity.

References

- Aristotle (1984). *The Complete Works*. Edited by J. Barnes. Oxford: Oxford University Press.
- Benacerraf, P. (1965). 'What numbers could not be'. *Philosophical Review* 74(1): 47–73.
- Burgess, J. P. (1999). 'Review of Stewart Shapiro's *Philosophy of Mathematics: Structure and Ontology*'. *Notre Dame Journal of Formal Logic* 40(2): 283–291.
- Burgess, J. P. (2014). 'On a derivation of the necessity of identity'. *Synthese* 191: 1567–1585.
- Enderton, H. (1977). *Elements of Set Theory*. San Diego, CA: Academic Press.
- Fraenkel, A., Bar-Hillel, Y. and Levy, A. (1973). *Foundations of Set Theory*. Amsterdam: Elsevier.
- Goodman, N. (1979). 'The problem of counterfactual conditionals'. In his *Fact, Fiction and Forecast*, Cambridge, MA: Harvard University Press, pp. 3–27.
- Hale, B. and Linnebo, Ø. (2020). 'Ontological categories and the problem of expressibility'. In B. Hale (ed.), *Essence and Existence; Selected Essays*. Oxford: Oxford University Press, pp. 73–103.
- Hale, B. and Wright, C. (2012). 'Horse sense'. *Journal of Philosophy* 109(1–2): 85–131.
- Hellman, G. (2001). 'Three varieties of mathematical structuralism'. *Philosophia Mathematica* 9(3): 184–211.
- Hewitt, S. T. (2012). 'Modalising plurals'. *Journal of Philosophical Logic* 41(5): 853–875.
- Hossack, H. (2000). 'Plurals and complexes'. *British Journal for the Philosophy of Science* 51: 411–443.
- Hossack, H. (2006). 'Reid and Brentano on Consciousness'. In M. Textor (ed.), *The Austrian Contribution to Analytic Philosophy*. London and New York: Routledge, pp. 1–36.
- Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number: What Numbers Are and How They Are Known*. London: Bloomsbury.
- Kripke, S. (1980). *Naming and Necessity*, Cambridge, MA: Harvard University Press.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
- Linnebo, Ø. (2010). 'Pluralities and sets'. *Journal of Philosophy* 107(3): 144–164.
- Linnebo, Ø. (2016). 'Plurals and modals'. *Canadian Journal of Philosophy* 46(4–5): 654–676.
- Marcus, R. B. [Ruth. C. Barcan] (1947). 'Identity of individuals in a strict functional calculus of second order'. *Journal of Symbolic Logic* 12: 12–15.
- Mendelson, E. (2015). *Introduction to Mathematical Logic*. New York: Chapman and Hall/CRC.
- Plato (1997). *Complete Works*. Edited by J. M. Cooper and D. S. Hutchinson. Indianapolis, IN: Hackett Publishing Company.
- Prior, A. (1955). *Formal Logic*. Oxford: Oxford University Press.
- Quine, W. V. (1953). 'Reference and modality'. In *From a Logical Point of View*. Cambridge, MA: Harvard University Press.
- Rumfitt, I. (2005). 'Plural terms: Another variety of reference?' In J. L. Bermudez (ed.), *Thought, Reference and Experience*. Oxford: Clarendon, pp. 84–123.
- Russell, B. (1911). 'On the relations between universals and particulars'. *Proceedings of the Aristotelian Society*, New Series, 12(1911–1912): 1–24.
- Russell, B. (1912). *The Problems of Philosophy*. New York: Henry Holt; London: Williams and Norgate.
- Russell, B. (1918). 'The philosophy of logical atomism [Lectures I and II]'. *The Monist* 28(4): 495–527.

- Russell, B. (1919a). 'The philosophy of logical atomism [Lectures III and IV]'. *The Monist* 29(1): 32–63.
- Russell, B. (1919b). 'On propositions: What they are and how they mean'. *Proceedings of the Aristotelian Society*. Supplementary Volumes 2: 1–43.
- Russell, B. (1919c). *Introduction to Mathematical Philosophy*. London: George Allen & Unwin.
- Russell, B. and Whitehead, A. N. (1910). *Principia Mathematica*. Vol. I. Cambridge: Cambridge University Press.
- Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press.
- Shapiro, S. (2008). 'Identity, indiscernibility, and ante rem structuralism: The tale of *i* and *-i*'. *Philosophia Mathematica* 16(3): 285–309.
- Uzquiano, G. (2011). 'Plural quantification and modality'. *Proceedings of the Aristotelian Society* 111(2–2): 219–250.
- Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.
- Wittgenstein, L. (2003). *Tractatus Logico-Philosophicus. Logisch-Philosophische Abhandlung*. Frankfurt a.M.: Suhrkamp.

A Summary of My Current Views

Keith Hossack

1. Relation dualism

Dualism is the doctrine that the mental is irreducible: the true theory of mind must include something mental amongst its primitives. This mental something need not be the soul: event dualism is the more modest theory that there are irreducibly mental events; still more modest is relation dualism, which only asserts the existence of a relation that is irreducibly mental. Relation dualism is a doctrine in the philosophy of mind, but it has interesting connections with epistemology and also with metaphysics.

2. Awareness as the irreducible mental relation

Relation dualism needs to identify which relation it takes to be fundamental in the theory of mind. Since mind is that which knows, relation dualists can seek the fundamental mental relation by an analysis of the concept of knowledge. If knowledge is warranted true belief, then since knowledge is factive, one's having warrant excludes the possibility of one's belief being false. The relation dualist should therefore take the fundamental mental relation to be a relation that warrants belief.

According to Brentano, every mental state presents an object, and this 'direction toward an object' is the mark of the mental. Should relation dualism take the fundamental mental relation to be Brentano's presentation relation? Surely not, for to do so would lead to scepticism about the external world. Brentano says the 'presented' object may possess only intentional inexistence – it need not exist in reality. If an external object that is presented might be mere appearance, its being presented does not exclude the possibility of one's belief being false.

According to the representational theory of mind, the propositional attitudes are the fundamental mental relations. On this theory, an experience as of an external object is a psychological relation to a proposition that represents a circumstance. But relation dualism should not take propositional attitudes to be the fundamental mental relations, for that would again open the door to scepticism. If the fundamental mental relations are relations to propositions, there is a veil of propositions between the mind and the world: a Cartesian deceiver could undetectably remove the external world without

disturbing the veil. So one's standing in a mental relation to a proposition that represents a circumstance does not exclude the possibility that the circumstance does not exist.

If we are to avoid scepticism, we must suppose that the fundamental mental relation is a relation to a *fact*, i.e. to the circumstance itself. If one stands in a mental relation to a circumstance – if one is in actual 'mental contact' with it – then one's mental state does exclude the possibility that the circumstance does not exist. So to avoid scepticism, relation dualism should postulate that there is a mental relation to a circumstance which warrants one's belief in the circumstance. Call this relation the *awareness relation*. There is then the following connection between knowledge and the awareness relation: S knows that A if S's true belief that A is warranted by S's awareness of the circumstance that A.

3. Circumstances

Taking the fundamental relation to be awareness allows relation dualism to avoid scepticism, for awareness is a relation between a mind and the circumstance itself, which foils the Cartesian deceiver. If the external world is removed, the circumstances it contained are removed too, so one would not stand in the same mental relations, so one's mental state would be different. Standing in the mental relations to the external world in which one actually stands excludes the possibility that the external world does not exist.

A truth, a true proposition, can be about a time, but is not itself in time. In contrast, a circumstance is a being that exists in time: a temporary circumstance exists only at a particular time, an eternal circumstance exists at every time. Temporary circumstances include events, states and processes; eternal circumstances are those whose existence can be known by logical, mathematical and ethical reasoning.

4. Epistemological ramifications of relation dualism

4.1 Consciousness

If S has a conscious experience, there is something it is like for S to have that experience. This property of the experience is called its phenomenal character. If a mental state has a phenomenal character, then necessarily the subject of the state is aware of its having that character. For example, a state could not be pain unless it was felt as pain. Conversely any state that is felt as pain really is pain – even 'phantom limb pain' really is pain. We can account for the necessary connection between having a pain and feeling the pain by supposing that the pain and the feeling of pain are one and the same identical event. This account, which I call the identity theory, entails that, necessarily, a mental state with phenomenal character occurs if and only if its subject is aware of its occurrence.

4.2 Faculties

Conscious states are self-announcing, since awareness of a conscious state is identical with the state itself. But most circumstances of which we are aware are known by

the operation of some faculty and are not self-announcing. For example, if a subject sees that the cat is on the mat, the circumstance of the cat's being on the mat must have caused the subject's faculty of vision to cause their awareness of the circumstance.

The operation of the human visual system gives rise to a visual experience and to awareness of the cat's being on the mat. How are the experience and the awareness related? Disjunctivists say that they are one and the same thing – the experience just is the subject's being visually aware of the cat on the mat. Disjunctivists say it is conceivable that there could be a perfect hallucination subjectively indistinguishable from the experience. Such a hallucination would not be awareness, so they say it would not be the same mental state as the veridical experience, although it would be subjectively indiscriminable. But it seems more natural to say that the veridical experience and the hallucinatory experience are identical mental states, since by hypothesis they have exactly the same functional role and exactly the same qualitative character. The experiences differ only in their effects: when veridical, the experience causes awareness, occurring as it does in an epistemically favourable context; when hallucinatory, the same experience does not cause awareness, because it occurs in an unfavourable context. Thus the relational state of the subject of the veridical experience is different from the relational state of the subject of the hallucinatory experience, although it is the same experience in both cases. This analysis allows the awareness theory to reject the sceptic's premiss that in the good and bad cases one's total mental state is the same.

4.3 Scepticism

Vision is a faculty, a power of the mind to acquire awareness of the visible circumstances. Similarly the other senses give us awareness of the circumstances proper to them. The sceptic may put forward an argument that we do not have such awareness. It is vain to attempt to argue with the sceptic, for an argument needs premisses, and whatever our premisses are, the sceptic will say that relying on such premisses begs the question. Therefore the correct response to scepticism is Moore's: any valid argument which entails that Moore doesn't know he has hands must have a false premiss, since the conclusion is ridiculous.

In addition to the senses, other human faculties include intelligence, which gives us scientific knowledge, empathy which gives us knowledge of other minds, and reason, which gives us *a priori* knowledge. Moore's way with scepticism allows us to affirm the existence of these other faculties also. Since it's ridiculous to suppose that Moore doesn't know that the sun will rise tomorrow, we must acknowledge the faculty of intelligence, which allows us to learn from experience. Since it's ridiculous to suppose that Moore doesn't know that other people have thoughts and feelings just as he does, we must acknowledge the faculty of empathy. And since it's ridiculous to suppose that Moore doesn't know that twice two are four, that no contradiction is true, and that some things are good and some actions right, we must acknowledge the faculty of reason.

5. Metaphysical ramifications

5.1 Circumstances

Relation dualism says there is an irreducibly mental relation, so relation dualism is false if no such entities as relations really exist. Bradley rejected the theory of relations on the grounds that it could not explain, supposing given some things and a relation, what it would be for the things actually to stand in the relation. Russell answered Bradley by appealing to the category of entities he called ‘complexes’. According to Russell, what it is for the things to stand in the relation is for there to exist the complex whose constituents are the things and the relation. If we interpret Russell’s ‘complexes’ as our circumstances, the reply to Bradley is that, for example, the cat and the mat stand in the *is-on* relation if there exists the circumstance of the cat’s being on the mat. Thus relation dualism must endorse a metaphysics of the real existence of circumstances. As we are taking awareness to be the fundamental mental relation, for each kind of knowledge that our faculties give us, we must expect there to be a corresponding kind of circumstance, our awareness of which warrants the knowledge.

5.2 Quantities

Events, states and processes are temporary circumstances, and all of them can be objects of visual awareness: one can see an event such as a flash, a process such as a wave, or a state such as the pavement’s being wet. But we can also see such beings as the ‘individual man or horse’, which according to Aristotle fall in the category of substance. So should the awareness theory allow that we can stand in the awareness relation to entities in the category of substance?

According to Aristotle’s hylomorphic theory, an individual substance comes into being when suitable matter receives the proper form. For example a table comes into being when suitable wood is arranged tablewise. Call a state of tablewise arrangement of some wood a ‘table-stage’. A plurality of such states, if properly connected both causally and spatiotemporally, may be said to comprise the life of a table: call such a plurality of stages ‘table-related’.

The awareness theory can say that one sees a table if one’s belief in the table’s presence is warranted by one’s visual awareness of the wood’s being arranged tablewise. But predicates such as ‘arranged tablewise’ are said of a quantity of matter, not of an individual, and predicates such as ‘table-related’ are said of a plurality of states and not of a single state. The subject term of such predications denotes not an individual but a quantity, so a language that can express such predications needs its quantifiers to range over quantities as well as individuals. In the predicate calculus, the quantifiers range only over individuals. A more general logic is therefore required, where the quantifiers range over quantities also. This more general logic is mereology, the pure *a priori* logic of quantity, as axiomatized by Tarski. The quantities for which Tarski gave axioms were continua – regions of three-dimensional Euclidean space. But his axioms apply equally to discrete quantities, where they emerge as the logic of plurals.

5.3 Mathematics

The logic of quantity allows us to attribute properties to quantities. One obvious property a quantity has is its size, and this leads us to the *a priori* science of mathematics. We can take a natural number to be the absolute size of a plurality, and a real number to be the relative size of a continuum. The axioms for the natural and real numbers, and for the ordinal numbers also, can then all be deduced from mereology plus the *a priori* axioms that Euclid calls the Common Notions.

5.4 Logically complex circumstances

The faculty of reason gives us *a priori* knowledge of mereology and indeed of logic generally. Since our logical beliefs are knowledge, they are warranted by our awareness of logical circumstances. For example, we know that everything is identical to itself, so there exists the circumstance of everything's being identical to itself. The logical properties and relations that are constituents of this circumstance include the logical property *all*, which a predicable has if each thing instantiates it, and the logical relation *identity*, in which each thing stands to itself.

Logically complex circumstances crop up also in the theory of causation. Circumstances are the causal relata: for example, the cat's being on the mat might cause the dog to bark. But some causes are logically complex, as when the stable door's not being locked causes the horse's escape. The logically complex circumstance of the stable door's not being locked has the constituent *not*, the logical relation in which a predicable and a subject stand, if the subject does not instantiate the predicable.

Hossack's bibliography

Books

Geometry and the Concept of Space, DPhil Thesis (Oxford 1986).

The Metaphysics of Knowledge (Oxford: Oxford University Press 2007).

Knowledge and the Philosophy of Number: What Numbers Are and How They Are Known (London: Bloomsbury Academic 2020).

Articles and chapters

A Problem about the Meaning of Intuitionist Negation, *Mind* 96/394 (1990): 207–219.

Access to Mathematical Objects, *Critica* 23/68 (1991): 157–181.

Constructivist Mathematics, Quantum Physics and Quantifiers, *Proceedings of the Aristotelian Society*, supplementary vol. 66/1 (1992): 61–97.

Intolerant Clones, *Mind*, 103/409 (1994): 55–58.

Plurals and Complexes, *The British Journal for the Philosophy of Science* 51/3 (2000): 411–443.

Self-Knowledge and Consciousness, *Proceedings of the Aristotelian Society* 102/1 (2002): 163–181.

- Consciousness in Act and Action, *Phenomenology and the Cognitive Sciences* 2 (2003): 187–203.
- Reid and Brentano on consciousness, in Markus Textor (ed.), *The Austrian Contribution to Analytic Philosophy* (London and New York: Routledge, 2006), pp. 1–36.
- Vagueness and personal identity, in Fraser MacBride (ed.), *Identity and Modality* (Oxford: Oxford University Press 2006), pp. 221–241.
- Seeing an individual, in MM McCabe and Mark Textor (eds.), *Perspectives on Perception* (Berlin: De Gruyter 2007), pp. 47–64.
- Actuality and Modal Rationalism, *Proceedings of the Aristotelian Society* 107/3 (2007): 433–456.
- Concept and content, in M. Stone (ed.), *Reason, Faith and History: Philosophical Essays for Paul Helm* (Farnham, UK: Ashgate 2008), pp. 205–218.
- Précis of *The Metaphysics of Knowledge*, *Dialectica* 65/1 (2011): 71–73.
- Replies to Comments, *Dialectica* 65/1 (2011): 125–151.
- A Correspondence Theory of Exemplification, *Axiomathes* 23 (2013): 365–380.
- Sets and Plural Comprehension, *The Journal of Philosophical Logic* 43/2–3 (2014): 517–539.

Book review

- Review of Stewart Shapiro's *Philosophy of Mathematics: Structure and Ontology*, *The Philosophical Quarterly*, 50/198 (January, 2000): 120–123.

Confronting Facts: On Hossack's *The Metaphysics of Knowledge*

Mark Sainsbury

1. The basicness of knowledge

In his dazzlingly bold and original *The Metaphysics of Knowledge* (2007), Keith Hossack suggests that whereas belief is a relation to a content, knowledge is a relation to a fact. Since facts and contents are radically different kinds of things, so are belief and knowledge. We should not try to analyse knowledge as a supplemented version of a state that occurs in mere belief. Rather, we can use knowledge to explain various other interesting epistemological notions: warrant, rationality, content, reliability, epistemic defeat and justification. It's a breathtaking project, undertaken with great energy and clarity.

A central idea that shapes the book is a version of the 'knowledge first' approach: a certain kind of knowledge is conceptually primitive and metaphysically fundamental. This basic kind is *knowledge-of*, which directly relates a subject to a fact, as contrasted with knowledge that something is the case, which involves a content as well as a fact. A fact is a structure of a universal and one or more particulars. Sameness of fact is just a matter of sameness of the universal and the particulars. Nothing intensional, like mode of presentation or content, is involved in knowledge-of facts.

Knowledge-that involves knowledge-of, but also involves a content: a mode of presentation of a fact. Hossack proposes the following logical form for 'S knows that A':

$$(x)(p)(x \text{ is a mental act} \wedge p \text{ is a fact} \wedge \text{content}(x) = \text{that-}A \wedge \text{that-}A \text{ is a mode of presentation of } p \wedge S \text{ knows of } p \text{ in virtue of } x)$$

(7 – all page references in this style are to Hossack 2007)

On this view, knowing *that* A is not simply being related to a fact, but is a complex state which obtains in virtue of relations to both a content, which is a mode of presentation of a fact, and also to the presented fact. Content is what makes 'propositional attitudes', like believing-that, non-extensional: the same fact may be presented in different ways, and a seeming presentation of a fact may not present a fact (as in the case of false belief). The mental act that witnesses the quantification in the displayed definition is a

belief whose content is *that-A*, so knowledge-that and believing-that both involve a relation to the same *kind* of thing (a content), and may relate to the same *specific* thing of that kind (*that-A*).

In Hossack's view, the ancient astronomers, before the familiar discovery, were aware of the fact that Hesperus is Phosphorus, for this is the fact that Hesperus is Hesperus, and logic alone acquainted them with the fact. Of course, were one to *say* that the ancients knew of the identity of Hesperus and Phosphorus, one might *implicate* the falsehood that they knew that Hesperus is Phosphorus. But this implicature is cancelable. In Hossack's example, such a cancelation is felicitous in 'Pharaoh was aware of the shining of Hesperus, though I do not imply that he was aware that it was Hesperus that was shining'. Assuming that 'aware of' is the same as 'knows of', and 'aware that' is the same as 'knows that', cancelation is rather harder to hear as felicitous for the case of identity. We would be disinclined to affirm 'Pharaoh knew of the identity of Hesperus and Phosphorus, though I do not mean to imply that he knew that it was Hesperus that was identical with Phosphorus'. English also has the locution 'ignorant of' and one might expect this to hold just when 'knows-of' does not. Yet it seems hard to resist the thought that Pharaoh was ignorant of the identity; that is, did not know of it.

However, I see no need for Hossack's position to reflect nuances of English idiom, for he has more significant fish to catch, notably the central claim that knowledge is conceptually and metaphysically more basic than belief. Conceptual basicness is straightforward, metaphysical basicness less so.

2. Conceptions of basicness

First, conceptual basicness. A natural definition is this: X is conceptually more basic than Y if X figures in the analysis of Y but not conversely. Hossack has more than one argument aimed at showing that the concept of knowledge is more conceptually basic than the concept of belief. One that especially struck me goes as follows. Belief is a relation to a content, and a true content is a mode of presentation of a fact. There can be different modes of presentation of the same fact, as we have just seen in the distinction between 'Hesperus is Hesperus' and 'Hesperus is Phosphorus'. How are modes of presentation, or more generally contents, to be individuated? Hossack suggests we should appeal to Evans' Fregean 'intuitive criterion of difference': contents are different if it is possible rationally to believe one but not the other (4). This obtains if there is no obvious rationally correct inference from one to the other, and 'an inference is rationally correct only if it is possible by its means to pass from knowledge of the premisses to knowledge of the conclusion'. Belief content is individuated by the intuitive criterion, which depends on the notion of rationality, and this in turn is explained in terms of knowledge. He concludes: 'it looks as if belief must be explained in terms of knowledge, and not the other way round' (5).

Hossack throws out this provocative argument rather briefly, and there are various possible responses. Perhaps content is not to be individuated by the intuitive criterion

of difference, but in some other way. For example, perhaps a content is a structure of concepts, where concepts are individuated by their origins and not by any semantic or epistemic features (as in Sainsbury and Tye 2012). Perhaps the notion of rationality is primitive, or can be explained in terms of truth and consistency rather than knowledge, a view possibly attributable to Davidson (1990). Hossack himself mentions other reasons to say that the concept of knowledge is conceptually fundamental, for example the long history of failures to provide a conceptual reduction. This is inconclusive in its nature, whereas the argument that starts with the intuitive criterion of difference, if it could be firmly established, would be conclusive.

This illustrates what conceptual fundamentality is: a relation among concepts fixed by some notion of analysis. It is harder to say what metaphysical fundamentality amounts to. Presumably it should be a relation among things that are not concepts: objects, properties or facts. The two relations would mirror one another if the following held: concept C1 is (conceptually) fundamental relative to concept C2 just if what C1 stands for is (metaphysically) fundamental relative to what C2 stands for. This would enable us to infer metaphysical fundamentality from conceptual fundamentality. But the equivalence is implausible in both directions. Colour concepts have a good claim to be (conceptually) fundamental relative to any other concepts, including concepts for wavelengths or surface reflectance profiles, since they are unanalysable; but presumably the colours themselves are not (metaphysically) fundamental relative to wavelengths, reflectances and perhaps perceptual systems. For failure in the other direction: an elementary particle is presumably (metaphysically) fundamental relative to any non-elementary particle or object, but the concept of being elementary is (arguably) analysable in terms of negative concepts like non-complexity or indivisibility, where these concepts are explained by their positive counterparts, as applied to non-elementary objects.

Even if we grant Hossack that knowledge is conceptually fundamental, it remains a further question whether it is metaphysically fundamental. The distinction is not salient in his work. For example, the book's main claim is at one point expressed:

Fact-knowledge is the relation that I take to be conceptually primitive and metaphysically fundamental. (8)

I regard these as two independent claims.

A positive suggestion is this: X is metaphysically more fundamental than Y if there could be X-worlds which are not Y-worlds, but no Y-worlds which are not X-worlds. It's not clear that Hossack holds that all knowledge is fundamental in this sense. A crucial feature of his account is that knowledge is caused by belief, under suitable conditions. An immediate thought is that this sounds as if it makes knowledge 'composite', in Williamson's sense (2006), which might appear inconsistent with being metaphysically fundamental relative to the components. Hossack denies this charge. Williamson's compositeness involves independent elements, for example an internal state like belief, and an external state, like knowledge-favourable circumstances, where the two elements are independent. By contrast, Hossack's view is that an internal state

of belief, along with a suitable context, jointly cause knowledge. 'Since joint causation has no simple logical analysis, we should not expect a simple analysis of how knowledge depends on intrinsic state plus context' (19).

This seems to be a claim about conceptual dependence, whereas I was hoping to find an argument for the metaphysical fundamentality of knowledge. The claim about causation, however, appears to make knowledge non-fundamental, since it is dependent on belief. This dependence is a claim he emphasizes:

We know in virtue of believing, because in a suitable context belief causes knowledge. (259)

If, as in the earlier displayed quotation (Hossack 2007:8), knowledge-of is metaphysically fundamental one would expect it to be possible for it to obtain in the absence of knowledge-that. But it's not clear that this really is possible. Hossack certainly doubts that it is actual, though he considers that an omniscient God might have no use for belief or perception, since 'he already has all the knowledge there is to be had' (267), and so might have knowledge-of without knowledge-that. But having no use for belief or perception does not ensure no knowledge-that. Omniscience would seem to require some knowledge-that. For example, God needs to know what knowledge-that his creatures possess, and *knowledge-of* the fact that X knows that *p* ensures *knowing-that p*. If God can hear our prayers, he needs to understand their content, which means knowing something like: they are praying to me that I will do such-and-such. It's hard to see how this could be captured in a purely knowledge-of way. It seems that, on Hossack's view, knowledge-of is possible only in situations in which there is knowledge-that, which makes knowledge-of non-fundamental in a straightforward sense.

Despite the claimed fundamentality of knowledge-of, we might wonder how we would be significantly worse off without it. On an opposed view, it's an artifact of our ascriptions of knowledge, rather than a specific kind of knowledge. We speak of knowledge-of when we wish to bracket content, and go straight to the referents of the content. But a full account of our mental operations requires specifying content. To describe someone in terms of their knowledge-of, rather than their knowledge-that, is to give a less informative account of their mental state. We say which facts they are related to by their knowledge, but fail to register how those facts are presented. If we were trying to describe the development of astronomy, it would not be helpful to say that everyone (with minimal logical acumen) always knew of the identity of Hesperus and Phosphorus. That would make it mysterious how their telescopes helped them make a discovery.

Hossack's apparent commitment to there being (perhaps only contingently) no knowledge-of without knowledge-that misses an opportunity: that of giving an account of a special feature of perceptual experience. Without any serious damage to the overall position, perceptual experience could be characterized as a distinctive source of knowledge-of which is not caused by knowledge-that. This gives knowledge-of a special metaphysical and explanatory role.

3. Inference and immediacy in testimony and perception

Hossack argues convincingly that we should not regard perceptual knowledge and testimonial knowledge as inferential, a respect in which the phenomena are alike. But he goes further, showing sympathy for the view that there is value in a notion of the ‘testimony of our senses’. The point about non-inferentiality seems to me correct, and Hossack supplies at least one novel argument for the conclusion, which I’ll shortly describe. There also seems to be an opportunity to use Hossack’s apparatus to explain a deep and important difference between testimony and perception.

Neither testimony nor perception is inferential, in the sense that neither testimonial nor perceptual knowledge needs to be derived from a premise affirming the obtaining of some knowledge-favourable state of the witness or the perception. For testimony, this kind of inferential knowledge would require it to be the case not merely that the witness is reliable or knowledgeable, but that the hearer of the testimony knows this. Yet examples make it plain that we can acquire knowledge by testimony without even having to reflect on whether the witness is reliable or knowledgeable. Similarly for perception: inferential knowledge would require not merely that the conditions are favourable for perception, but that we know this. Hossack gives an ingenious argument to show that we can attain perceptual knowledge while being unable to know that our senses are reliable. Even if we lack the fine discrimination that would enable us to know that our situation is (by a hair’s breadth) one in which our senses are reliable, that they are in fact reliable is enough to permit perceptual knowledge (240–241).

Yet, in at least one critical respect, testimony and perception are very different. In testimony, something does stand between us and the fact we may come to know, namely the testimony itself. We have to ‘read’ it (in Hossack’s phrase: i.e. extract its informational content). We first have to figure out what the witness testified. Only once we have done this do we have a chance to know, not merely that the witness testified that *p*, but that *p*. The testimony is a ‘veil of witnessing’.

One can agree with Hossack that perception is not inferential without agreeing that non-inferentiality is enough to make it ‘immediate’:

a visible fact causes one’s faculty of vision to cause a visual experience which if the context is suitable causes one immediately to know the visible fact. (244)

Agreed, the process is not one of inference. But on this account the process is causal: something (a visual experience) stands between the known fact and the knowledge of the fact. In testimony, that intermediate element is the testimony itself, which is straightforwardly a mediator. In favourable cases, the testimony is caused by the fact testified to, and in turn causes knowledge of that fact in the audience. For Hossack, perception is similarly structured: something stands between the visible fact and the knowledge of it, namely, perceptual experience. This mediating causal role, linking fact to knowledge, is appropriate for testimony. But it can be doubted whether it does full justice to the nature of perception.

Hossack allows that the analogy between the two ways of gaining knowledge is not perfect, and he expresses it in a nicely turned phrase:

There is a contrast here with perception: one may already know one faces a blue cube, but vision will continue to tell one this for as long as one continues to look ...; but speakers who keep telling conversational partners what they already know are soon partnerless. (247)

True enough: perception is insistent in a way that testimony is not. Indeed, it can be insistent even when the subject knows quite well that it is illusory, as when a knowledgeable subject sees a Müller–Lyer drawing: there is no repressing perception's insistence that the lines are unequal. But the most striking respect in which perception differs from testimony is its 'here and now' quality, or 'presentness'; its capacity to 'disclose' or 'directly present' how things are, enabling 'direct confrontation' with reality. Perception is more like a window than a witness. In perception we confront the world; in testimony, just testimony. These observations, however correct, are merely metaphorical. It seems to me that Hossack's framework can be used to move beyond metaphor to something more literal and precise.

4. The openness of perception

There are many ways of attributing perceptual experiences, but one that seems specially apt to evoke the openness we are trying to explain treats tropes as objects of experience (along, possibly, with individual objects). An example by Mark Johnston¹ can be used to support this idea. You are looking at your car's red leather in a light too dim for you to see the redness. The leather looks red to you because some non-perceptual, but potentially reliable process (e.g. memory), 'fills in' what's missing. On the basis of this perception, you might come to know that the leather is red. But it remains that you do not see the redness of the leather, because the lighting is too poor. The crucial point is the possible truth of both of these attributions:

You saw that the leather was red.

You did not see the redness of the leather.

Seeing a trope (a property instance) demands a relation to that trope which may not be required for the truth of related constructions taking sentential complements of perceptual verbs.

There are less fanciful examples of this contrast. You might see that the knapsack she was carrying was heavy, without seeing its heaviness, for heaviness is not as such visible. You might see that she was likely to cross the road, without seeing her likeliness to cross the road. There are also converse differences: you might hear the last note without hearing that it was the last note, or feel the touch of a mosquito without feeling that a mosquito was touching you.

We can apply the contrast to a case already mentioned, suggesting that the first of these could be false even when the second is true:

Pharaoh saw that Phosphorus shone at dawn yesterday.

Pharaoh saw the shining of Phosphorus at dawn yesterday.

The first, using the seeing-that construction, requires Pharaoh to know that the star he saw at dawn was Phosphorus. The second, using a sentence nominalization referring to a trope, does not. The first is nonextensional, the second is extensional.² The first characterizes perception-based knowledge; the second characterizes the perception itself. The first leaves the openness of perception a mystery; the second helps us understand what it involves.

Words for tropes sometimes figure in what Hossack calls 'sentence nominalizations':

if A is the sentence ' a is F at t ', then its nominalization A^* will be 'the F -ness of a at t ' or ' a 's being F at t '. . . whenever we have ' S knows that A ' we shall also have ' S knows of A^* ', where A^* is the nominalization of A . (8–9)

This seems to be in conflict with Johnston's claim. 'The redness of the leather' is a nominalization of 'the leather is red'. So the inference pattern in the displayed quotation licenses the inference from 'You saw that the leather is red' to 'You saw the redness of the leather'. But I am persuaded that this is the wrong verdict. I think Hossack could abandon the inferential principle just displayed without any damage to the overall picture. Then he could use the contrast between knowledge-of and knowledge-that to help explain the distinctive features of perception.

'Sees that p ' entails 'knows that p ', so it can be regarded as a determinate of the determinable knowledge-that. Perceiving a trope entails gaining knowledge-of that trope, but does not ensure any specific knowledge-that. Attributions of seeing tropes are extensional, like attributions of knowledge-of. These connections between perceptual verbs and knowledge verbs allows us to use essentially Hossack's apparatus:

to perceive is to gain knowledge-of by means of the senses.

Perceiving may result in knowledge-that, if a cognitive process initiated by the knowledge-of selects one of possibly many ways of conceptualizing that input to deliver knowledge-that. This is not an inferential process, for inference requires knowledge of premises, which is knowledge-that; and this process starts just with knowledge-of. No trope-related *premise* need be known by the subject, even if a trope is known. The initial input states are properly described extensionally, the output states are only adequately described non-extensionally.

Although perception and testimony are similar in not requiring inference, they differ in how the knowledge-of/knowledge-that contrast applies. Testimony's starting point is knowledge-that or belief-that, and this may lead to knowledge-of. Perception's starting point is knowledge-of, and this may lead to knowledge-that.

The proposal is that the openness of perception should be explained in terms of its taking tropes as among its objects, generating extensional, and in this sense 'direct', knowledge-of. What is normally referred to as perceptual knowledge, knowledge-that, is a conceptualized product of this knowledge-of. The notion of 'openness' with which we began, and allied notions like 'directness', 'disclosure', 'presentness', 'here-and-nowness' are largely metaphorical. The present claim is that we can use knowledge-of to supply a more literal basis for them.

We can start by seeing if the Hossackian picture I have proposed can explain some more literally expressed contrasts between perception and testimony:

1. Perceptual information is limited to the here-and-now: it can deliver knowledge-of only of what is currently perceptible. By contrast, there are no limits to the kind of information testimony can supply, or to what time the testimony relates.
2. Perception has a vividness that testimony can never attain. That's because it's knowledge-of, unmediated by the concepts involved in knowledge-that.
3. Perception is typically much richer than testimony. You cannot begin to describe what you see in full detail: your testimony relating to an experience would greatly underdescribe what you experienced. That's because knowledge-of is independent of the conceptual resources or conceptual deployment needed for knowledge-that.

To amplify the last point: if there is motion in the scene, it will normally be much too fast for you to deliver descriptions rapidly enough to keep up. Your colour and shape vocabularies are inadequate to do justice to the visual scene. 'Green' is hardly specific enough, and suggests a uniformity that is rare in nature. 'Leaf-shaped' is manifestly inadequate since leaves come in a huge variety of different shapes. This is perhaps why philosophers try to keep to absurdly simplified examples, like seeing a bulgy red tomato, or a blue cube. But our knowledge-of tropes, we learn from perception, are not constrained in these ways; which also explains how creatures for whom testimony is unavailable may be acute perceivers.

4. Perception has a coercive quality that testimony lacks. With our eyes open, we cannot help seeing. With our ears uncovered, we cannot help hearing. We see and hear the world itself, not just testaments about it. We can ignore a testimony we know to be a lie. We hear the words but have no inclination to believe them. An experience we know to be illusory, like the apparently moving scrolling messages on some public notice boards, is different. We have to engage in active resistance in order to avoid forming the corresponding belief. It's hard not to see the letters as really moving, even when we know the effect is produced by the successive illumination of stationary bulbs. A partial explanation is that our concepts are sometimes under voluntary control: we have some choice about which concepts to exercise. Our confrontation with tropes is not in the same way under our control: if we are normally sighted, and our eyes are open, we cannot but see what is before us.

These explanations compare favourably with another familiar explanatory proposal: in perception, the objects or scene perceived are said to be 'constitutive' of the experience.

This does not give us any insight into the window-like character, or openness, of perception. If anything is constitutive of experience, its subject is, but even if subjects are always aware of themselves in perception (which seems doubtful) their awareness of themselves is completely unlike their awareness of the perceived scene. Contrast with a summary of the Hossackian position I am proposing: perception's openness consists in being knowledge-of, the feature that makes it capable of being reported extensionally, independently of concepts.

Consider the following three sources of empirical knowledge: (i) perception; (ii) explanations, predictions and theories based on what is perceived (as in empirical science); and (iii) testimony. For the second two categories, knowing-that or believing-that precedes knowledge-of. Perception is distinctive because in that case knowledge-of comes first, and knowledge-that is derived from it. Perception is open: it delivers in the first instance knowledge-of; and this in turn often produces, but is not mediated by, knowledge-that or belief-that.

5. Hallucination

In hallucination, nothing is perceived, so the experience provides no knowledge-of tropes. We can make sense of a 'perfect' hallucination, an experience in which nothing available to consciousness marks it as a hallucination. A perfect hallucination has the same phenomenology as a possible veridical experience. Hence, it seems, the shared phenomenology cannot be 'explained' in terms of 'knowledge-of'; moreover, knowledge-of cannot correctly identify the phenomenology in the veridical case, since it would be incapable of revealing that case as one whose phenomenology is also available in hallucination.

This is a familiar form of objection, one that has led to so-called 'disjunctivist' approaches; and that is the kind of approach that seems natural in the present context. As an initial response, explanations do not always need to supply necessary conditions: the window broke because a brick was thrown at it, providing (a salient part of) a sufficient condition but not a necessary one; she succeeded in voting because she mailed in a ballot paper on time (though she could also have voted in person). A necessary condition may be a disjunction of sufficient conditions.

Knowledge-of provides an illuminating sufficient condition for perceptual phenomenology in veridical cases. This is not undermined by the fact that the same phenomenology may arise in a different way. There is also a closer connection between the cases: in perfect hallucination, it *seems to the subject as if* she has knowledge-of. So knowledge-of does in the end contribute to the explanation of the phenomenology in both cases: in veridical cases, its presence suffices for the openness of perception; in hallucinatory cases, the seeming openness consists in our seeming to have knowledge-of.³

This last point can be made stronger by the observation that both hallucinatory and veridical experiences are produced by a system whose teleological function is to produce veridical ones: a perceptual system is a heritable capacity that supplies knowledge-of an organism's environment often enough to promote the organism's

survival. This explains why creatures endowed with perceptual abilities came to outperform those lacking the abilities, thus ensuring the survival of the ability over generations. Hallucinatory experiences resemble veridical ones in being defective products of the same system.⁴ Unlike knowledge-that, perceptual knowledge-of does not require the exercise of concepts, and so is available to creatures who lack them.

What is supposed to happen does happen in successful cases, but not in unsuccessful ones. This is just what one should expect, for one would not expect phenomenology as such to have an externalist feature⁵ like necessarily involving knowledge-of. To summarize in a distinctively disjunctivist way, highlighting the role of knowledge-of:

to perceive, or to seem to perceive, is to be in a state produced by a system whose function is to enable subjects who possess it to acquire knowledge-of the subject's environment.

We are intuitively inclined to describe perception in metaphorical terms: it is *direct, open, here-and-now, disclosive, immediate, vivid, rich and coercive*. It is not that the ascriptions are incorrect, but that they are merely metaphorical and not deeply informative. The Hossackian position changes that. Instead of these unclear ascriptions, it makes central use of a clear, non-metaphorical and revealing notion: acquiring knowledge-of the world by means of the senses.

Notes

- 1 Johnston (2006): 278. Revisiting his text, I realize my exegesis is open to question.
- 2 'If fact-knowledge is a simple relation between a mind and a fact, then it should be possible to report fact-knowledge in a way that is referentially transparent' (8).
- 3 A full account would also discuss illusion. That would require a careful account of what illusion is, saying enough to determine whether seemingly converging railroad tracks, tilted coins and so on should properly be counted as sources of illusions.
- 4 Compare Johnston (2004): 151: 'Hallucination is a degenerate state, a failure of the visual system to function properly'.
- 5 Though for other kinds of externalist features (notably historical ones) this expectation has been claimed by many to be erroneous. See for example Lycan (2001), Tye (2009): §8.6.

References

- Davidson, D. (1990). 'The structure and content of truth'. *Journal of Philosophy* 87: 279–328.
- Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Johnston, M. (2004). 'The obscure object of hallucination'. *Philosophical Studies* 120: 113–183.
- Johnston, M. (2006). 'The function of sensory awareness'. In T. S. Gendler and J. Hawthorne (eds), *Perceptual Experience*. Oxford: Oxford University Press, pp. 260–290.

- Lycan, W. G. (2001). 'The case for phenomenal externalism.' *Philosophical Perspectives* 15: 17–35.
- Sainsbury, M. and Tye, M. (2012). *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*. Oxford: Oxford University Press.
- Tye, M. (2009). *Consciousness Revisited. Materialism without Phenomenal Concepts*. Cambridge, MA: MIT Press.
- Williamson, T. (2006). 'Can cognition be factorized into internal and external components?' In R. Stainton (ed.), *Contemporary Debates in Cognitive Science*. Oxford: Blackwell, pp. 291–306.

Who Knows?

Mary Margaret McCabe

1. Plato's forms and the objects of knowledge

I have had the good fortune to talk to Keith over the years about philosophy of all sorts, and especially about Platonic idealism. Keith is much taken with the theory of forms, especially with its connections to mathematical realism.¹ And indeed there are places in the dialogues – notably, one might think, the central books of the *Republic* – where the ties between metaphysics and mathematics are strong indeed.² Here Plato's account of knowledge may seem completely dominated by the realist thought that knowing just is having some kind of access or connection to these ultra-real things, forms. Now I don't dispute that there are some ultra-real things lurking in the central books of the *Republic* and elsewhere in the dialogues;³ nor that – in some sense or another – they are taken to be objects of knowledge. There is a question of scope here, which I shall not discuss further: do the ultra-real objects exhaust the objects of knowledge or is there knowledge also of other things? Leaving that question aside, we still should be wary of allowing a general thought about the objects of knowledge (so, a general strong realism) to slip into a different one altogether: that (on Plato's view, or anyway) knowledge *just is* some kind of direct contact with or grasp of or affection by its objects and that we have given an adequate account of Platonic knowledge exactly when we have given an account of these objects. Call this different thought the *object-exhaustive* view of knowledge.

The object-exhaustive view of knowledge has some undesirable consequences for the account of the philosopher, and especially the philosopher-ruler, in the *Republic*: and these consequences, I believe, obscure the significance of Plato's account of knowledge and its role in epistemic virtue. Consider three.

First, for object-exhaustiveness, knowledge will occur exactly when (and just because) some object of knowledge (some individual form) turns up epistemically for the knower. This epistemic turning-up might occur in two different ways: either all the objects turn up at once (knowledge is 'all-or-nothing'⁴) or they turn up piecemeal (form by form, perhaps), knowledge lapsing when its object is no longer there. Knowledge, on this view, would be like some thin versions of perception.⁵ This brings out the central feature of object-exhaustiveness, that it tells us about objects, but little

about the knowing subject; if the objects explain knowledge exhaustively, the knowing subject seems to be *passive* to them.⁶

Second, if object-exhaustive knowledge is piecemeal, while this may be congenial enough to many contemporary accounts of knowledge (where the paradigm case of knowing is often 'knowing that p'),⁷ it makes nothing of the suggestion, expressed frequently in the *Republic* and elsewhere, that knowledge is systematic like, perhaps, a science.⁸ That systematicity, in turn, may mirror a view of knowledge as a developed disposition of the knower, even a virtue (such as wisdom).⁹ Any suggestion that knowledge just happens through a causal relation between reality (piecemeal or systematic) and some mind that just happens to be in its way seems, however, incompatible with the thought that knowledge is a virtue: virtue is not by happenstance.¹⁰

Third, the object-exhaustive view garbles the political project of the *Republic*, in which the philosopher is, *by virtue of her wisdom*, somehow or other good at ruling.¹¹ If there is some correlation between wisdom and knowledge,¹² that *she is a knower* has some kind of explanatory role. That explanatory role needs to tell us about the knower herself, not just about her passive encounter with knowledge's objects. Indeed, she is picked out as a privileged and virtuous individual, because she is the *subject* of knowledge, which explains her claims to know, and thence her claims to rule.¹³ What is more, the *Republic* asks just why knowledge (laboriously acquired as it is over years of hard work and abstaining from any fun) is a good in itself *for her*:¹⁴ and why (which may or may not be the same question) knowledge is either itself (a) virtue or the foundation for the virtue of wisdom or in fact exactly what wisdom is.

2. Epistemic pluralism?

From this last thought, and en route to thinking about the knowing subject, we might ask too about epistemic pluralism.¹⁵ Should we think of knowledge as coming in many distinct kinds and fashions, or are the different modes in which we speak of knowledge all somehow connected, or are they all, in fact, different ways of talking about the same thing? This may be a question we especially want to ask of Plato, who uses an extravagant lexicon for the terms of knowledge (words we could gloss variously as: knowledge-that; knowledge-how; knowledge-what; skill, art or craft;¹⁶ science; understanding; reason; intuition – and wisdom¹⁷).

However, the Platonic lexicon of a different, but connected, semantic group is narrow. For when he talks about knowing, he also talks about *saying*: not only about how we articulate our knowledge,¹⁸ but about how we acquire or progress in it,¹⁹ how it is explanatory or accountable,²⁰ and how we ourselves give accounts of it to others.²¹ These different features of our epistemic activities all turn, in Plato's lexicon, on a limited set of cognate terms, the verb for saying (*legein*) and its nominal correlate, what is said (*logos*). So when he talks about words, statements, assertions, reasons, accounts, explanations, speeches, agreeing and disagreeing, contradicting and explaining, the range of words that he uses is parsimonious. Although in carefully demarcated passages

he will use the vocabulary of causation, explanation, blame and responsibility (the group of *aitios* words²²) his overwhelming preference is to use words derived from the *legein* root. Indeed, contrary to his variations with the knowledge lexicon,²³ he uses *legein* cognates repeatedly in the same passage, shifting between verb and noun, and adding prefixes from time to time, but repeating the same vocabulary.²⁴

This is so, and marked, especially in passages where the arguments are *about* saying, or asserting, or whatever, and are commented upon in the frame dialogue in exactly the same terms.²⁵ For, recall that the Platonic dialogues vary in their narrative style: some are framed by a narrator (or a series of narrators);²⁶ some are directly in dialogue.²⁷ In the former case, it is commonplace for the frame to report ‘he said’, ‘he asked’, ‘he replied’ and so forth.²⁸ In such cases, the frame dialogue sometimes comments on what is said in the reported discussion, and sometimes does so by reflecting on the terms in which, or the descriptions under which, something or other is said.²⁹ Even in dialogues like the *Theaetetus* where both the outer frame and the main dialogue are in direct speech, there is a framing effect when the protagonists comment on what is said by themselves and also by imagined others;³⁰ so that there is reportage from one part of the dialogue to another. The dialogue form has thus a varied but thoroughgoing second-order character, often brought out by the ways that the vocabulary of saying carries over from the reported to the reporting, and is reflected upon there.

In the *Theaetetus*, notably, the topic is knowledge; and the topic is both the question at hand and part of the language in which it is questioned, as we shall see. But this goes hand in hand with the theme of saying. For one of the most promising attempts to define knowledge proposes that knowledge is true belief with a *logos*, a ‘saying’.³¹ This definition turns on how we should understand whatever may be meant by a *logos*: it may seem promising because even although this definition of knowledge runs aground in the dialogue itself, a different interpretation of true belief with a *logos* may after all succeed. Contemporary takes have been swift to read the proposal as ‘justified true belief’ (so that a ‘saying’ is a justification).³² I shall suggest that deeper scrutiny of this running set of connections between knowing and *logos*, knowing and saying, especially in the context of dialogue, may offer different illumination, especially when it comes to thinking about the knowing and saying *subject*.

3. The *Meno* question

Platonic epistemology has enjoyed quite a vogue, recently,³³ not so much the realist epistemology of the *Republic* and its form of the good, as the puzzle we find in a dialogue with a more contemporary feel – the *Meno*. Late in the *Meno*, Socrates and Meno confront a (now) famous puzzle:

Socrates ... Suppose someone who knew (*eidôs*) the way to Larissa, or wherever you like, went to Larissa and led others there too, surely he would lead them correctly and well?

Meno Certainly.

Soc What about this: suppose someone who had the right belief about the road to Larissa but had never been there nor knew (*epistamenos*) the way – suppose this person went there and led others too: surely he too would lead them correctly?

M Certainly.

Soc So as long as this person had the correct belief about the things of which the other had knowledge (*epistêmê*), he will be no worse a leader, believing true things but not knowing (*phronôn*) them, than the one who knows (*phronôn*) it?

M No worse.

Soc So true belief is no worse a leader than wisdom (*phronêsis*) when it comes to correctness in action. And this is what we left out just now when we were considering what virtue is, when we suggested that wisdom (*phronêsis*) alone leads correct action; for true belief does so too.

M It seems so.

Soc So correct belief is no less useful than knowledge (*epistêmê*)?

M Well, Socrates, but only to this extent – the person who has knowledge (*epistêmê*) will always succeed, while the person with correct belief will sometimes succeed and sometimes fail.

Soc Oh, really? Surely the person who always has correct belief will always hit the mark, so long as he believes correctly?

M It seems that this must be right; in that case, then, Socrates, I wonder why knowledge (*epistêmê*) is much more valued than correct belief, and even why they are distinct.

Meno 97a–c³⁴

Stuck on the road to Larissa, we have little reason to prefer the guide who knows over the guide who has (what we know to be? *ex hypothesi*?) true belief.³⁵ Indeed – suppose we were assured that our guide *always* has true beliefs in all the relevant possible circumstances,³⁶ it might be much more efficient to engage a guide who has not needed laboriously to acquire knowledge. Why would we have any reason to prefer the person who knows? What *more is there* (if anything) to knowing than regular (always) true believing? And how does that ‘more’ explain why we hold knowledge in such high esteem? So there is a pair of questions here: what is the *relation* between knowledge and true belief; and what is the *value* of knowledge, over and above regular true belief?³⁷ Socrates’ gnomic answer – that knowledge is tied down by the ‘account of the reasons’ (*aitias logismos*)³⁸ – may be little immediate help. But many have thought that the *Theaetetus* – itself a dialogue in conversation with the *Meno*³⁹ – may provide illumination.

4. Piecemeal and systematic: the question of addition

The *Theaetetus* seems to pastiche the so-called ‘Socratic’ dialogues of definition. Rather than seeking to define some virtue or another or even virtue in general, it seems to ask instead about the epistemic correlate of our having such a definition: ‘what is knowledge?’ or, ‘how do we know what knowledge is?’ Of course, the pastiche may make us think that after all knowledge is a virtue: it may even do so by rejecting the route of epistemic pluralism:

Soc ... The question, Theaetetus, was not this one: of what things there is knowledge? Nor was it how many kinds of knowledge there are; we didn’t ask the question because we wanted to count them, but to understand knowledge itself, what it is.

Theaetetus 146e

In the third part of the *Theaetetus*, after two unsuccessful attempts, Socrates proposes a third definition of knowledge: that knowledge is true belief with a *logos* (201c ff.). What would it be to have a *logos*? Socrates puts to flight pretty smartly the first construal of the proposal, that knowledge is ‘making one’s own thought clear through speech with verbs and nouns, just as in a mirror or in water, making an image of the belief in the stream of words’ (206d). For if knowledge were (something like) true belief + an utterance (the stream of words), that would give us no account of what would make knowledge so distinctive (it would get nowhere with Meno’s questions, either). Second, perhaps knowledge enumerates the elements of something (so a *logos* is some kind of exhaustive list)? Socrates objects that this can be poorly grounded (some of the enumerated elements may be present in ways vulnerable to mistake). Third, perhaps true belief needs to give an account of the distinctiveness of its object; but this seems circular, since grasping the object at all seems to be by virtue of its distinctness. So the proposal to add a *logos* to a true belief to produce knowledge gets us nowhere after all. Has something important been missed?

Maybe we have simply got hold of the wrong version of a *logos*, and the wrong idea about saying. Perhaps, first, we have got ourselves wrongly hung up on something propositional: the grammatical model of the ‘stream of words’ leads us astray at the outset. The dialogue does focus on a specialized set of sayings, namely definitions (such as would answer the question ‘what is knowledge?’). But definitions need not have a fixed grammar nor even be propositional in form.⁴⁰ We may as well talk about knowing the aardvark as about knowing what an aardvark is as about knowing that the aardvark is a member of the only surviving species (*Orycteropus*) of the family of *Orycteropodidae*⁴¹ – without finding ourselves switching into a different *kind* of knowledge each time. Instead the distinctiveness of the objects of knowledge (definitions) will embrace the plurality of modes of knowledge. Just this pluralism is captured by the pluralism of a *logos*: a saying, or an utterance, or a list, a reason, an account and so forth. But (to remind ourselves of Meno’s other question) this pair of pluralisms – knowledge and

logos – gets us nowhere with explaining why knowledge should be special, and valuable, compared with true belief. So perhaps we need to base knowledge on some special kind of *logos* that also tells us about its value?⁴²

The *Theaetetus* does not contemplate the possibility that knowing and saying are special when they are about forms (although some think that this is an omission in the dialogue that we are supposed to notice and fill in for ourselves⁴³). But it does hint at some way of managing a grand kind of *logos* – something like an ‘account’ – compatible with the view that knowledge should not be lumpy or piecemeal but systematic. Socrates’ ‘dream’ in the *Theaetetus* (201e–202c) allows for a complex structure of both reality and accountability in which the elements of things can only be perceived and named but still form part of a complex of things by virtue of ‘common’ expressions (the expressions of formal relations between the elements – relations of identity and difference and various kinds of ordering) and so constitute a *logos*. And then it becomes possible to ‘give and receive a *logos*’ (202c), a procedure which is essential for knowledge.⁴⁴

On such a model, the object of knowledge might be thought of as a domain, specialized by picking out the elements (are they essences?) within that domain, and correspondingly expressed or described by a complex and systematic account. This, indeed, may be where the *Sophist* takes the dream: to an account of *kinds* (are they essences?) that is described as dialectic:

Eleatic Stranger Well then – since we agreed that kinds undergo mixing with each other in the same way [sc. as the notes in music and the letters in writing], surely it must be with some kind of knowledge proceeding through *logoi* (accounts?) that someone would have to show us correctly which kinds harmonise with which and which kinds resist each other? And further, doesn’t he have to know whether there are any kinds that run through all of them and hold them together so that they are capable of mixing together, and again, when there are divisions, whether other kinds running through wholes are causes of the division?

Theaetetus How could this not need knowledge – probably just about the greatest . . . ?

Stranger . . . Shall we say that to divide according to kinds and not to confuse one kind for another or the other for it, that this is dialectical knowledge?

Sophist 253c–e

So if the domain of kinds gives us the objects of knowledge, it is important that it is structured: and that structure is articulated as a *logos*, perhaps an account, something we can give and receive in a dialectical context. The *Theaetetus* baulks at the dream; but perhaps by the end of the dialogue we are supposed still to ask whether this is the sort of account that could be *added* to a true belief to make it knowledge.⁴⁵

Mere addition seems to miss some of the point, not least in supposing that for both epistemic states, there is underlying belief. For the contrast between some true belief and the knowledge of a domain seems thoroughgoing.

First, we cannot turn *any old* true belief into some knowledge just by adding some suitable things, whether they be *logoi* or warrants or justifications. *Not every* true belief

can be transformed into knowledge by some additions unless there is nothing special at all about knowledge's domain.⁴⁶ But equally, if philosophers are to rule, as the *Republic* proposes, their epistemic attitude when they deal with the particular world should not turn into a belief just when they do, or else they have no epistemic advantage after all (call this the 'philosopher-ruler problem').⁴⁷

Second, if knowledge is of a domain, not piecemeal, then the having of some single true belief, even if it concerns that domain, could not, on its own, count as knowledge.⁴⁸ On what we might think of as an *additive* view, knowledge might be in part constituted by beliefs, with whatever appendages fix them as knowledge. On a *counter-additive* view, belief and knowledge are quite different in structure and in function: so they are different in kind, and neither constitutes the other. This does not make their contents exclusive (the content of some beliefs may turn up in the domain of knowledge) but their conditions are quite distinct.⁴⁹

Third – to return to Socrates' closing impasse – the kind of account (*logos*) we find in the case of knowledge (the account of a whole domain) works quite differently for the content of a belief: there, the *logos* may be exhausted by a single statement, or proposition, or whatever. So belief and knowledge will be quite different cognitive modes, as is appropriate to epistemic pluralism. Does this mean that the associated sayings, the *logoi*, are pluralist too, equivocal as they line up with different epistemic modes (e.g. articulating a simple belief versus giving and receiving a dialectical account)?⁵⁰ And does this take us any further forward in accounting for the value of knowledge?

5. Priorities

On the object-exhaustive view of knowledge, even if there is a difference between the objects of belief (piecemeal objects, perhaps a broader range of objects than for knowledge) and the objects of knowledge (domains), the difference between knowledge and beliefs seems merely to be one of scope; belief may have more particular content than knowledge while knowledge may be harder to come by than belief. However, a remarkable passage at the centre of the *Republic* insists on thinking about these two epistemic features as different, and distinct, *powers or faculties* (*dunameis*). Here the process that has someone come to know or understand changes their state of mind⁵¹ from what is called *doxa* (belief) to what is called *epistēmē* (knowledge). For these are quite different 'powers', *dunameis*.⁵² In part their content may remain the same: whatever we need from the philosopher-ruler, we at least need her knowledge brought to bear on the ordinary issues of political life (the philosopher-ruler problem). But the epistemic characterization is quite different just because the state of mind, the power which enables us to do whatever we do, is different.⁵³ The reference to the knowing subject and her capacities is ineliminable.

How should we capture this feature of knowledge? A domain of objects can be thought of as knowledge or science; but a faculty or a capacity needs stronger talk. Here, knowledge is closer to what we think of as *understanding*, itself a state of mind that is distinct from belief (supporting the counter-additive view). Understanding, as

an epistemic state, is far more domain-orientated than belief (which is readily piecemeal); but to understand understanding demands too that more be said about the possessor of the state of mind. If we focus on a contrast between knowledge and belief as states of mind, they do indeed seem distinct, instantiated separately, as the *Republic* demands.

Consider first a difference in the causal histories of the two epistemic states. If understanding and its power are fixed in part by a complete domain, they will differ formally from how we might understand the cognitive state of belief. The beliefs I may have, true or false, may turn up piecemeal, depending on my local and occurrent experiences (consider the very first example of a belief in the *Republic*, Polemarchus' belief that Socrates and Glaucon are about to go back to the city from the Piraeus, just as he first encounters them, 327c). That my epistemic state is *belief* is, in the least case, just the result of that feature of my occurrent beliefs: the instances come first, and they may not build up into anything more systematic easily, or ever.⁵⁴ In the case of understanding, however, this seems to be in reverse. Understanding on this view has, as its objective correlate, an interconnected system of something or other (whether kinds, or essences, in the Platonic mode, or bodies of science, or whole societies or whatever). Without that system, understanding has not yet been achieved, and the system is what explains that understanding is what it is. But that is not enough to explain what we mean here when we say that she understands; the luminous state of her mind needs more. When we speak of knowledge, we may abstract away from its possessor: we can describe a body of knowledge and its parts without reference to any person at all. For understanding, this is not so; even when understanding is discussed in the abstract, it makes a nonsense of it to imagine it simply in the domain of objects. This state of mind, when it occurs, determines that some particular saying or thought is an instance of understanding, rather than of belief. When I understand something, that is when I have *understanding* as my epistemic state: and the reference to the subject is ineliminable. For individual instances, that state of mind, understanding, comes first in the order of explanation. At the same time, it is – as the *Republic* explores – last in the order of development: in the individual, understanding comes very late, if it comes at all.⁵⁵

But if understanding is first in the order of explanation, its relation to belief is not explained by an additive model. Furthermore, if understanding is last in the order of development, it may occur just when belief is left behind (understanding, on this account, might arrive all at once). But we might still wonder just what that involves. For the *content* of understanding will still include *some* of the objective content of belief (under penalty of the philosopher-ruler problem). So what are we to say, then, to differentiate the two epistemic states such that addition is not the account to prefer? The question has a sequel: if we think that somehow (as the third part of the *Theaetetus* suggests) there is a significant relation between understanding and whatever we mean by *logos*, what are we to say about *logos* such that it is not banal (not merely a statement of what is understood), but central to our understanding of understanding (the dialectical ability to give and take a *logos*), and without the risk of a regress (*logoi logôn*). How are we to understand the integrated relations of *logoi* such that we end up with the right body of objects and the understanding of it too?

6. The phenomenology of understanding

The *Theaetetus* poses its question in an oddly reflexive way, ‘we wanted to *know* what knowledge itself is’ or ‘we wanted to understand what understanding is’. For not only does it play on the contrast between this question, ‘what is understanding?’ and the familiar Socratic questions, ‘what is piety?’, ‘what is courage?’, but the language of the question – as Socrates takes pains to observe – leaches into the question itself: not just ‘what is understanding?’, but ‘how do we understand what understanding is?’. Equally, if we find an answer to the question about the essence of understanding, it will itself be an object of understanding.⁵⁶ The domains of understanding may be iterated and ordered, so that the domain is heavily structured, thoroughly three-dimensional, or even more. And it might be either obvious, or stipulated, that relations between items in the domain, between essences or kinds or whatever, are explanatory relations (this would echo the ‘account-giving’ of the *Meno*). So there is plenty to say about knowledge as understanding, already. But does the heavy structuring of understanding explain to us what counts as *our* having understanding, or as *our* having knowledge glossed in that way? Do these heavy structures explain the correlative state of mind? If *I seek to understand* what understanding is, am I satisfied when the objective domain of understanding contains an additional element that is the understanding what understanding is?

Suppose that we understand just if we somehow or other *have* those explanations in their domains. But what, I wonder, lurks in that ‘something or other’? What is it to *have* understanding, or to *get* an explanation? *How* – we might ask – does Plato think that understanding, accounted for in terms of its objective domain, is the epistemic state or property of a subject with a power or a capacity, the subject who understands the domain and understands too, perhaps, what understanding is? If I know that the essence of understanding is thus and so, what does the essence of understanding tell me about my own cognitive state when I have it? (And we might still wonder, what advantage, if any, does that cognitive state have over regular true belief?) Equally, when my understanding involves some kind of *logos*, what is there to that *logos* over and above its being incorporated into the domain in question? What is it for *me* to have such a *logos*, or to get it – or even, in parallel to the *Theaetetus*’ question, to *say* it?

Well, we had better *not* think that I have this epistemic state just by virtue of some passive *encounter* with the objects of the domain (any more than that I can say the *logos* just by virtue of coming across it lying around in my head⁵⁷). Of course that thought would run directly into the philosopher-ruler problem. And even Thrasymachus dislikes the idea that knowledge is somehow a passive affection . . .

‘How am I to persuade you?’ he [Thrasymachus] asked. ‘If you’re not convinced by what I said now, what more can I do for you? Do you want me to sit here and cram the argument in with a spoon?’

Republic 345b⁵⁸

. . . as, too, does Socrates:

how excellent it would be, Agathon, if wisdom were such as to flow from the fuller of us to the more vacant, if we touch each other, just as water always flows along a piece of material from a full cup into an empty one. If wisdom were like this, I would greatly value lying beside you, for I think that I would be filled with so much fine wisdom flowing from you. My own wisdom is a poor thing, contested and dreamlike, but yours is shining and great with promise.

Symposium 175d–e

And indeed – if understanding comes first in the order of explanation but last in the order of development – the question is not about how individual pieces of understanding get in my head; but how my head develops the power to apprehend them *as understanding*.⁵⁹ If we reject the additive model, this question becomes more urgent: for then we need to figure out just how the alterations in my epistemic state can be made, if that state is not merely constituted by the propositions that are common to the domain of understanding and belief along with whatever the add-ons would have been. Understanding will not be like Lego, but different in kind from its antecedents. How?

Perhaps the place to begin is with the phenomenology of understanding. In English, at least, it has both objective and subjective [I mean by this, ‘pertaining to the subject’] features. Objectively, we may think about understanding as directed at a domain, rather than a particular proposition (‘she understands the aardvark’ depends on the assumption that she understands a whole lot of other stuff about the aardvark beyond either the collocation of the letters of its name or the hairy quality of its nose); that it improves on rote-learning by knowing that the causal relations are indeed causal relations; and that it is factive: knowledge implies the truth of its content, and the falsity of its content denies the knowledge.⁶⁰ Subjectively, it has a vigorous personal reference (‘Oh, I understand!’; ‘Oh, I see!’ where the attention to the pronoun matters). It is thus reflexive (I report my own state of mind when I exclaim in this way) and can be reflective (I can explicitly record my own state of mind). It also carries a pleased implication of success: my state of mind is somehow luminous to me.⁶¹ The three features of the objective account of understanding are echoed, I propose, by the three features of the subjective one: so the objective domain is parallel to the personal claim; the causal structure to the reflexive and reflective report; and the factiveness to the implication of success. This rich complexity, I think, Plato saw.⁶²

7. The development of understanding

Central to Plato’s understanding of understanding are his repeated accounts of the *development* of understanding in its subject. This is, of course, a major theme of the *Republic*,⁶³ but it turns up elsewhere too. Consider the pairs of sophistic arguments that appear in both the *Theaetetus* (188a–189b) and the *Euthydemus* (276a–277d, and again at 283b ff.) to suggest that whether we think of the subjects of knowledge or the objects of knowledge, there is no possibility of either mistake or of epistemic change.⁶⁴ In response to the impasse produced by the sophists’ twin pairs of arguments in the

Euthydemus, Socrates proffers a contrast between ‘learning’ and ‘understanding’ and with it, I suggest, some account of how we might deal with the ways in which understanding comes first in explanation, and last in development.

First, as Prodicus says, you should learn the correctness of names. This is what our visitors are showing you, that you don’t understand that learning is the name men use for cases when someone from the beginning has no knowledge about some matter, and then later gets knowledge of it; but they use the same name for cases when someone already has the knowledge, and with this same knowledge considers that very same matter either in action or in saying. They more often call the latter understanding than learning, but they sometimes call it learning, too. But you had forgotten this, as they have demonstrated, that the same name is used for people in quite opposite conditions, for someone who knows, and for someone who does not. Pretty much the same thing was going on in the second question, too, when they asked you whether men learn what they know or what they do not know.

Euthydemus 277e5–278b2

His interest here may not be, as many suppose, in diagnosing an ambiguity in the words we use for learning and knowing (and so espousing some kind of epistemic pluralism), but rather in showing us something about aspect-change. For here the same verb (in this case ‘learning’) may be used for different stages of the same end-related process; and the end is described by the same term (‘learnedness’, perhaps). The sophists insist that learning can cause death (if when we learn we change and become someone else 283a ff.) or immortality (if our knowledge state is present, how can it ever have been absent? 296c-d); but Socrates insists that epistemic change and aspect-difference is centrally a matter of the state of the *subject*. The subjective phenomenon ‘I understand!’ is marked in Greek by the verb which we also use for the process by which we arrive at that understanding: I am learning (I learn) when I seek to understand, I have learned (I learn) when I succeed in understanding. This feature of subjective understanding works for Greek, then, as it does for English.⁶⁵

But how does the learning process produce the characteristic relation between the subject and the objective body of knowledge that we call understanding? If Socrates’ response to the sophists is right, then talk of understanding needs to include the subjective conditions, or the conditions for developing the subjective epistemic state, as well as their objective correlate. It is, then, no accident that Socrates discusses learning in connection with the state of the knower and with wisdom: the virtue of understanding.

Knowledge – on Plato’s account – is not simply poured in from the outside. Instead, there are internal conditions on its development: notably some cognition of one’s own epistemic state, some thought about its value and standing, and reflection on how it may be improved (there is a direct parallel here with the development of ethical virtue).⁶⁶ Consider, for example, the reflexive features of the Socratic *elenchus* (where part of the process is the compelling sense of puzzlement, as the interlocutor realizes that he doesn’t know what he thought he did).⁶⁷ Or notice the ways in which the changing view of the philosopher as she ascends from the cave is marked by her own

backwards looks and her awareness of her own blindness – and also by the ways in which the story of the cave is itself framed by a kind of reflective vision, as we the readers are asked to see (a verb often translated as to ‘picture’) the prisoners in the cave, ‘like ourselves’ (*Republic* 514a ff.). For it is on the back of these reflections that explicit and planned development takes place: they provoke further inquiry, or affirm progress made.

This reflection forms part of the machinery of psychological or epistemic change. It is not the case, of course, that knowledge or understanding requires reflection for *each* move or *each* proposition; rather, the process of developing the cognitive condition, the power, of understanding, demands at some stage that we ask the sorts of questions that the *Theaetetus* asks: how is it that we know what knowledge is? And here and now, do we know it? And what do we say of our understanding? So the subjective elements of knowledge or understanding – the subject’s epistemic state – will be structured in complex ways that condition successful development, just as are the objective elements. On this account, if knowledge is plausibly taken to be understanding, its conditions cannot be merely additions to true belief. Instead, knowledge should be understood in teleological terms: in terms of the ways in which it develops towards the end of understanding.

It is, therefore, no surprise at all that the *Theaetetus* begins its question with a reflexive condition, and continues in the framed dialogue with a detailed account of the ways in which we may describe the development of understanding (notably in the digression, itself an account of the teleology of epistemic development).⁶⁸ It is equally no surprise that the framing of the *Theaetetus* is an elaborately constructed direct dialogue within a dialogue, thematizing the theme of saying and how saying is reported. Where the surprise turns up, instead, is in the thin account of knowledge that is canvassed in the subsequent main arguments, and its equally thin account of *logos*, ‘saying’. For throughout the framed discussion both of the thesis that knowledge is perception and of the thesis that knowledge is true belief with a *logos*, the condition of the subject of knowledge is barely addressed. The same is true of the subject of *logos*: the privileged *logoi* that are canvassed as additions to true belief in the closing puzzles are understood simply in terms of their content, and not at all in terms of the background assumptions or the epistemic condition of the speaker. But then the closing impasse – and the implicit question about addition – asks just how far the objective domain provides the content of speech; and if it does, how the connections between one *logos* and another are made, if not by the speaker herself. What are we to say about *knowing*, and about *saying*, over and above what is known, and what is said?

8. Knowing and saying

There is a long story to be told here. Both the *Euthydemus* and the *Theaetetus* pose a series of sophistic puzzles about the relations between different *logoi*. The *Theaetetus*’ discussion of false belief shows that we have no good account to give of the mental condition of the person who delivers themselves of false beliefs or false statements; but that such an account is vital if we are to explain how we are ever wrong.⁶⁹ The

Euthydemus supports the impossibility of false saying by a hopelessly reductive account of change and identity (283d–e) alongside a challenge to the possibility of dialogue and disagreement (283e–286c). The connections made in both dialogues between questions of metaphysical continuity and of the persistence of the subject point towards a replacement of the reductive view by a sophisticated teleology: what the subject aimed to do, how the subject aspires and changes and develops over time and how her practices of saying are connected and explained. Thus, it is clearly a condition on knowledge, for Plato, that it is somehow connected to what the subject *does* with *logoi* of some richer kind than even propositions about essence: for it is a part of the process of learning, thinking and understanding that the subject can give and take accounts, *logoi*.⁷⁰ So it is part of Plato's rejection of the views of sophists about truth that saying, seeking the truth, giving an account, proffering an explanation, being receptive to the views of others, having a conversation – activities that are all semantically connected – are teleologically connected too. The parsimonious vocabulary of *saying*, of *logos*, shows up Plato's commitment to the thought that what we say and how we say is a part of an interconnected process in the saying subject, the process towards telling the truth and being accountable in the ways demanded by the *Meno* – regularly and with understanding.

The parsimony of the lexicon of *logos* lies at the heart of Plato's explanation of knowledge. I suggested that he uses an extravagant vocabulary for knowledge; and suggested that this should make us ask whether that vocabulary indicates a radical epistemic pluralism: does Plato think that there are all sorts of epistemic states which are only vaguely connected with each other, and in fact reflect different kinds of epistemic interest (knowledge-that, knowledge-how, understanding, science, etc.)? If his account of knowledge is based on his account of saying, we may answer this question in the negative. For a teleological account of knowledge – one, that is to say, which sees knowledge as the outcome of a long process of learning and acquired skill and sustained reflection – allows for some pluralism in the course of the process, without supposing that the different kinds of epistemic state are somehow radically distinct or disconnected. Instead, they are integrated, on this view, on a continuous and complicated process of acquiring understanding; integrated, that is to say, by the continuity of the accountable subject herself. If Plato eschews the additive account of the relation between true belief and knowledge, and if he focuses on understanding, rather than different models of knowledge, he can explain the development of the cognitive subject – in the *Republic*, for sure, but also in the framing discussion of the *Theaetetus* and elsewhere. They illustrate vividly how for Plato we can reasonably interchange the notions of knowledge and of understanding with that of wisdom: epistemic pluralism is well connected after all.

This leads us back to the second *Meno* question, of the value of knowledge. We might think that any response to the question why knowledge is more valuable than true belief will be framed (whether affirmatively or otherwise) in terms of its instrumental value – either will be worth having for the sake of our consequent success in our projects and in our lives overall.⁷¹ But the suggestion that knowledge *is* understanding or that it *is* wisdom prompts a quite different view of its value. Imagine a moment of saying, in surprised pleasure as someone tries to explain wave-particle

duality to you, ‘oh I understand!’ That eventual moment of understanding is not instrumental, at all. On the contrary, it is something that we find to be good in itself (suspend here on just how we should explain the metaphysics of value here – for my money, intrinsic in some of the sense explored by Korsgaard and Williams: that is, good in at least the sense of the origin or source of goodness⁷²). This intellectual value is, after all, a familiar one to those of us who cheerfully plough the back-furrows of the history of philosophy.⁷³

It tells us, I think, two things. The first is that understanding does not behave in the ways that contemporary epistemology expects knowledge to behave (it is not a response to scepticism, or even a way of figuring out something about some bizarre agricultural edifices⁷⁴). Rather, here the state of mind, arrived at in the ways that we acquire virtue (by habituation and reflection), is the source of its individual instances and the source of some at least of their value. The second is that what we do every day (in the back-furrows of ancient philosophy, at least) gives a sense to this kind of value. The value of understanding must be value in itself: it would be a mistake to hesitate to claim it.⁷⁵

Notes

- 1 *Knowledge and the Philosophy of Number* (Bloomsbury 2020). It is an honour and a pleasure to offer this discussion of Plato to Keith, to thank him for thirty years of brilliant friendship.
- 2 Starting with the geometry of the construction of the line (509d ff.), then its discussion of mathematical objects and practices (510bff.), leading into the role of mathematics in the education of the guardians (522bff.) [I use the page and line numbers of Slings’ text of the *Republic*]. See e.g. Vlastos (1988), Burnyeat (2000), Broadie (forthcoming).
- 3 E.g. *Republic* 507b ff., *Phaedo* 78b ff.
- 4 On this in the puzzles about recollection see Fine (2003 esp. ch. 2), Scott (1995, Part 1).
- 5 I reject this view of Plato’s account of perception at McCabe (2016).
- 6 The object-exhaustive view is, you may complain, a straw man: most commentators seek to downplay exhaustiveness, by suggesting that the grasp of forms (or other ultra-real objects of knowledge) is somehow more complicated than it is passive.
- 7 E.g. in the opening paragraphs of Goldman (1976).
- 8 E.g. *Republic* 510b–511d.
- 9 See, for example, *Euthydemus* 278e–282d. Epistemic virtue is much discussed these days: see e.g. Sosa (1980), Zagzebski (1996), Brady and Pritchard (2006), Pritchard (2007). In general, these views of epistemic virtue take it to be *analogous* to ethical virtue although e.g. Zagzebski’s view is stronger (that is especially true, perhaps, of the strand of virtue epistemology often described as virtue responsibilism, in contrast to e.g. Sosa’s reliabilism: see Wright (2018)). In Plato’s case the ethical and the epistemic are intimately connected, in ways that I begin to explore here. His argument is a long and complex one, but it begins with the closing stages of the *Meno*. As will become clear, I hope, he does not thereby subscribe to a view that gives knowledge instrumental value (or instrumental value alone) nor to a thin intellectualism.
- 10 Becoming virtuous is not done by rote learning: and this, indeed, seems to be a thought central to the puzzles about becoming virtuous in Plato, from the *Meno*

onwards; and in Aristotle, likewise, see e.g. *Ethica Nicomachea* II and elsewhere. I have benefited from discussion of this issue with Margaret Hampson.

- 11 Is the philosopher-ruler a 'she'? There is (given the arguments of *Republic* 451c ff.) no reason to deny that she could be; but it is arguable whether these arguments are themselves especially palatable to a feminist view. See here Woolf (forthcoming) and McCabe (2020).
- 12 To the relation between the two I return below.
- 13 See the discussion at *Republic* 518d ff. of something like the unity of the virtues; and the supposition that requiring the philosophers to rule is 'asking just things of just persons', 520e.
- 14 Glaucon's original challenge assumes that wisdom (or thought? *phronein*) is a good 'itself by itself', 357c: the *Republic* as a whole is designed to show that we may say the same of justice.
- 15 By this I do not mean what I take Boghossian to mean when he discusses epistemic pluralism, namely the possibility of plural epistemic systems (Boghossian 2006), but rather that there are all sorts of different epistemic businesses expressed in our knowledge-talk.
- 16 Depending on our translation of *technê*: compare, for example, *Charmides* 165b or (what may be a more searching exposé) *Euthydemus* 278e–281e. On skills and virtue see Annas (2011).
- 17 Just in the *Republic*, for example, *technê* (e.g. 341d etc.), *phronêsis* (505c, 621c), *sophia* (443e, 504a), *epistêmê* (443c, 478a), *noêsis* (511d, 523a ff.), *nous* (431c, 508c), *gnôsis* (476c, 508e), *theôria* (517d), *dianoia* (476b, 511a) *mathêsis* and associated verbs (511a–b). A regular association of knowing with seeing (McCabe 2016) runs parallel with an association between knowing and conversation, eventually finessing conversation as 'dialectic' (McCabe 2015, ch. 6): notice the *logos/legein* root of the expression *dialektikê*, it is connected, somehow or another, with saying or talking (see 511b ff., 525d–e).
- 18 E.g. 476a–b, 527a.
- 19 527a, 531e.
- 20 Successfully or otherwise, see e.g. 431a. See e.g. the *sullogizein* verbs e.g. at 517c, 519d.
- 21 531e, 533c.
- 22 Notably at *Meno* 98a and markedly at a central point in the *Republic*, 508a ff.
- 23 E.g. *Euthydemus* 278e–282d; or between the frame and the first sophistic episode of the same dialogue, 276a–278e.
- 24 So including the compound verbs *dialegesthai* and *sullogizein* as well as the uncompounded forms, see above nn.17 and 20.
- 25 See the extended argument about false saying at the centre of the *Euthydemus*, 283a–288d.
- 26 The *Republic*, for example, or the *Protagoras*, in both of which cases the narrator is Socrates, a fact played on within the dialogues, e.g. *Protagoras* 334c–d, *Republic* 393b–c.
- 27 *Gorgias*, or *Meno*.
- 28 Although of course the *Theaetetus* comments on this and leaves 'he said' and so forth out, 143c.
- 29 Trivially, 'did he really say that?', exemplified in the breaking of the direct frame into the indirect reported framed dialogue at *Euthydemus* 290e.
- 30 E.g. the 'dialogue' with *Protagoras*' head, 171c–d.
- 31 201d ff.

- 32 See e.g. a 'standard narrative' at McGlynn (2014) ch. 1.
- 33 E.g. Sosa (1980); Kvanvig (2003) ch. 1; Hyman (2011).
- 34 My translations throughout.
- 35 In discussing *doxa* Plato switches between the particular cognitive event – 'a belief', 'a knowledge' [for this see *Phaedo* 73c–d] – and the epistemic state – 'belief', 'knowledge' [for this see e.g. *Republic* 476 ff.]. It is important that this switch is visible in English; translating *doxa* as 'judgement' precludes it, so I do not.
- 36 My thanks to Nils Kürbis for this qualification.
- 37 Pritchard (2007). Compare and contrast the so-called 'swamping' problem for the value of reliabilism and elsewhere: the value of true belief swamps that of true belief + reliable means, just because the thing is instrumental from the start.
- 38 Grube translates 'by (giving) an account of the reason why'. The translation of this expression is contentious, of course: I shall wonder in what follows about Grube's parenthesis '(giving)' which does not in fact translate the Greek. 'Reason' here, *logismos*, derives from the *legein/logos* root.
- 39 There is a high level of intertextuality between *Meno*, *Euthydemus*, *Republic* and *Theaetetus*.
- 40 As Nehamas has persuasively argued (1999).
- 41 Itself the only member of the order of Tubulidentata.
- 42 Nehamas (1999).
- 43 See Burnyeat's magisterial analysis (1990).
- 44 By the end of this dense passage, of course, we are challenged to offer a univocal translation of the *logos* that will be given and received such that it constitutes knowledge. The density of 'saying' words through this passage is remarkable: moving backwards and forwards from frame to framed (e.g. in the frame at 201d4, 5, 6 followed by framed 201e2ff.).
- 45 Note the language of addition e.g. at 206c4, 208e4. Broadie (2016) is characteristically lucid on the strategies here.
- 46 If knowledge requires some grip on essences, for example, not all beliefs involve essences except at some very remote level (unless we suppose that all general predicates, for example, refer to essences).
- 47 There is a larger question here, for a different occasion, about whether knowledge is inevitably general, where belief is often particular; this is connected, of course, to the philosopher-ruler problem.
- 48 See here Pritchard (2007).
- 49 For consideration of this view of Plato see especially Harte (2017) and also Vogt (2012).
- 50 A theory of knowledge, focused on a domain and on how that domain is articulated, would move away from any interest in essences *singularim* or definitions likewise. This may explain the *Theaetetus*' mannered shift from asking about piety or courage to an apparently similar question about knowledge.
- 51 These are perhaps confusingly described as affections in the soul, *pathēmata en tēi psuchēi* e.g. at 511d7.
- 52 Harte (2017).
- 53 Here I sidestep the long debate about the scope and content of knowledge and belief in the *Republic* and about the soundness of this conclusion; but see Fine (2003) chs. 3 and 4.
- 54 This is implicit in the discussions of early childhood in books 2–3.

- 55 This explains philosopher rulers and also the two blindnesses of the cave (515d, 517a). On the risk of pessimism in this epistemology see Woolf (2004).
- 56 Nehamas (1999).
- 57 The *Theaetetus* puzzles over this kind of model of the mind, in the model of false belief that is the aviary. If grabbing the wrong bird is what it is to make a mistake, how can I explain my decision to grab the wrong bird? (Notice the elliptical language of taking, having and losing throughout, notably at 197c–d).
- 58 Here translation by Griffith, just because it is so neat.
- 59 The luminosity of subjective understanding may show that there is no risk here of a regress of knowing that I know.
- 60 This feature of knowledge is contested, but I have not space here to explore it further (see e.g. Buckwalter 2014). The factiveness of knowledge/understanding (if there is any) means not only that a true knowledge claim implies the truth of its content (if I know that p, then p) but also that if the content is false, the knowledge claim is falsified too. In this respect knowledge notably differs from general belief (although not, of course, from true belief). It has a bearing on the ways in which knowledge develops, by reflection on where it fails.
- 61 Williamson (2000), e.g. pp. 13–15.
- 62 Consider e.g. 508d.
- 63 The luminosity thought runs in tandem with the regular analogy of seeing to knowing.
- 64 I have discussed this *Euthydemus* material extensively in McCabe 2021.
- 65 E.g. *Theaetetus* 174b6; *Phaedo* 117c1; *Republic* 372e2; *Meno* 84d4: these examples are present tense, but perfective aspect. In each case, the ‘understand’ is seen as the answer to an ongoing series of questions or a train of thought or an argument.
- 66 So Glaucon’s challenge, 357c, treats knowledge as a suitable analogue for justice in the question of whether justice is a good in itself.
- 67 See e.g. *Gorgias* 489a–b; or Socrates’ account of his inquiries in Athens, *Apology* 21b ff.
- 68 172c ff.
- 69 Notably at issue in the aviary: 197c ff.: note the focus in the entire image on the states of the subject (‘having,’ ‘possessing’ etc.).
- 70 This is especially a feature of the *Republic*’s account of knowledge and dialectic: 531e, 533c.
- 71 The *Protagoras* might give us one such view, 347 ff.
- 72 Korsgaard (1982), Williams (2006).
- 73 Compare here Cassam’s account of ‘inquiry epistemology’ which has built in, I think, the value of inquiry in itself (see Cassam 2016).
- 74 The ‘fake-barns’ literature begins with Goldman (1976).
- 75 My thanks here to many people for conversations in different places: to John Cooper, Michael Coxhead, David Galloway, Verity Harte, Nils Kürbis and, of course, to Keith.

References

- Annas, J. (2011). *Intelligent Virtue*. Oxford: Oxford University Press.
- Boghossian, P. (2006). *Fear of Knowledge*. Oxford: Oxford University Press.
- Brady, M. and Pritchard, D. (2006). ‘Epistemic virtue and virtue epistemology’. *Philosophical Studies* 130: 1–8.

- Broadie, S. (2016). 'The knowledge unacknowledged in the *Theaetetus*', *Oxford Studies in Ancient Philosophy* 51: 87–117.
- Broadie, S. (forthcoming). 'Why the rulers' mathematical education?' In M. M. McCabe and S. Trépanier (eds), *Re-Reading the Republic*. Edinburgh: Edinburgh University Press.
- Buckwalter, W. (2014). 'Factive verbs and protagonist projection', *Episteme* 11(4): 391–409.
- Burnyeat, M. (1990). *The Theaetetus of Plato* with translation by M. J. Levett. Indianapolis and Cambridge: Hackett.
- Burnyeat, M. (2000). 'Plato on why mathematics is good for the soul'. In T. Smiley (ed.), *Mathematics and Necessity: Essays in the History of Philosophy*. Oxford and New York: Oxford University Press, pp. 1–81.
- Cassam, Q. (2016). 'Vice epistemology', *The Monist* 99: 159–180.
- Fine, G. (2003). *Plato on Knowledge and Forms: Selected Essays*. Oxford: Clarendon Press.
- Goldman, A. I. (1976). 'Discrimination and perceptual knowledge', *The Journal of Philosophy* 73: 771–791.
- Harte, V. (2017). 'Knowing and believing in *Republic* 5', In V. Harte and R. Woolf (eds), *Rereading Ancient Philosophy: Old Chestnuts and Sacred Cows*. Cambridge: Cambridge University Press, ch. 7.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number*. London: Bloomsbury.
- Hyman, J. (2011). 'The road to Larissa', In M. de Gaynesford (ed.), *Agents and Their Actions*. Oxford: Wiley Blackwell, ch. 3.
- Korsgaard, C. (1983). 'Two distinctions in goodness', *Philosophical Review* 92: 169–195.
- Kvanvig, J. L. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- McCabe, MM (2015). *Platonic Conversations*. Oxford: Oxford University Press.
- McCabe, MM (2016). 'The unity of virtue', *Supplementary proceedings of the Aristotelian Society*.
- McCabe, MM (2020). 'Philosopher queens; the wrong question at the wrong time?' In E. Vintiadis (ed.), *Philosophy by Women*. London: Routledge, ch. 13.
- McCabe, MM (2021). "'Who's who and what's what?' A response to commentators on 'First chop your logos'". *Australasian Philosophical Review*, <https://doi.org/10.1080/24740500.2020.1716668>.
- McGlynn, A. (2014) *Knowledge First*. New York: Palgrave.
- Nehamas, A. (1999). 'Epistēmē and logos in Plato's later thought', In A. Nehamas, *Virtues of Authenticity*. Princeton, NJ: Princeton University Press, ch. 11.
- Pritchard, D. (2007). 'Recent work on epistemic value', *American Philosophical Quarterly* 44(2): 85–110.
- Scott, D. (1995). *Recollection and Experience*. Cambridge: Cambridge University Press.
- Sosa, E. (1980). 'The raft and the pyramid: Coherence versus foundations in the theory of knowledge', *Midwest Studies in Philosophy* 5(1): 3–26.
- Vlastos, G. (1988). 'Elenchus and mathematics: A turning-point in Plato's philosophical development', *American Journal of Philology* 109(3): 362–396.
- Vogt, K. (2012). *Belief and Truth: A Skeptic Reading of Plato*. Oxford: Oxford University Press.
- Williams, B. (2006). 'Plato's construction of intrinsic goodness', In B. Williams, *The Sense of the Past*, ed. M. F. Burnyeat. Princeton, NJ: Princeton University Press, pp. 118–137.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Woolf, R. (2004). 'The practice of a philosopher', *Oxford Studies in Ancient Philosophy* 26: 97–129.

- Woolf, R. (forthcoming). 'Dolphins and dialectic'. In MM McCabe and S. Trépanier (eds), *Re-Reading the Republic*. Edinburgh: University of Edinburgh Press.
- Wright, S. (2018). 'Virtue responsibilism'. In N. E. Snow (ed.), *The Oxford Handbook of Virtue*. Oxford: Oxford University Press.
- Zagzebski, L. (1996). *Virtues of the Mind*. Cambridge: Cambridge University Press.

Knowledge-first Epistemology and the Input Problem¹

Scott Sturgeon

1. Introduction

My first job in Philosophy was at King's College, London. As soon as I arrived there I began talking philosophy with Keith Hossack. Those conversations occurred over a number of years and I learned a terrific amount from them. I am grateful to Keith for his willingness to explain things which I do not understand, and for his great skill at doing so. It is a real pleasure to contribute to a volume in his honour.

Keith's work promotes knowledge-first epistemology, and his primary method of doing so is to show how knowledge can be used to explain things in metaphysics. Knowledge-first epistemology can also be promoted, of course, by showing how knowledge can be used to explain things in epistemology. But when that happens – in the most natural way, anyway – one immediately faces what I call the *Input Problem*. We'll see that this problem is faced by other epistemologies. But knowledge-first epistemology faces it in a particularly clear way. So that will be our focus.

The chapter unfolds as follows. In §§1–6, I explain the major components of a full-dress theory of epistemic rationality. In §§7–8, I explain how knowledge-first epistemology slots into that theory to generate the Input Problem. In §§9–13, I discuss what to say about the problem.

2. The theory of epistemic rationality

The full-dress theory of epistemic rationality is composed of four major theoretical parts:

- (i) a theory of epistemic targets,
- (ii) a theory of evidential information,
- (iii) a theory of evidential support, and
- (iv) a conversion theory.

The theory of epistemic targets pins down foci of epistemic evaluation. This bit of full dress theory specifies the sorts of things subject to epistemic evaluation: states, events, persons, movements-of-mind, and so on. The theory of targets is often presupposed in epistemic literature, but full-dress theory will make it explicit. Such a theory will involve a transparent listing of things subject to epistemic evaluation.

The theory of evidential information specifies which pieces of information count as evidence for a situated agent. This will involve a story about why some but not all pieces of information fall into an agent's total body of evidence, why they qualify to be pieces of evidence for the agent, and so on. Like the theory of targets, the theory of evidential information is often presupposed in epistemic literature, but it too will be explicit in full-dress theory. Such a theory will contain a transparent story about how a situated agent comes to possess particular bodies of evidence.

The theory of evidential support then specifies how pieces of evidential information support further pieces of information. In turn facts about this help to pin down how, exactly, an agent should exploit evidence when updating opinion. The theory of evidential support systematizes facts about evidential support, detailing large-scale structural patterns in the way support plays out, as well as local support relations between contents. The systematization of it all is what makes formal epistemology possible. The theory of evidential support is where that systemization happens.

Once the theory of targets, the theory of evidential information, and the theory of evidential support are in place, the conversion theory kicks in. This is a story about how agents should configure their epistemic targets in light of their overall situation. And it too is often left tacit in the literature, for epistemologists often (wrongly) assume that an agent should simply bring her epistemic targets into alignment with the upshot of her evidential situation, and her evidential situation is fully specified by the theory of evidential information. Our discussion of the Input Problem will make clear why this is not the case.

The full-dress theory of epistemic rationality, then, consists in a theory of targets, a theory of evidential information, a theory of evidential support, and a story about how the upshot of these things convert into epistemic prescription. Our next task is to see how these bits of theory fit together. This can be easily done by examining the choice-points for elements of full-dress theory. Consider each element in turn.

3. Epistemic targets

The main division in the theory of epistemic targets is between work on coarse-grained and work on fine-grained targets. The former develops theory for rational belief, disbelief and suspended judgement. The latter does so for rational confidence: credence, interval-valued confidence, and so on. Debate between coarse- and fine-grained theorists is a hot-button issue right now in both epistemology and the philosophy of mind. In the next section, we'll see one way that knowledge-first epistemology can take a stand on a coarse-grained state, knowledge, which drives the rationality of an agent's epistemic targets.

4. Evidential information

Not only does full-dress epistemic theory require an approach to epistemic targets, it also requires an account of evidential information. This bit of theory specifies which information counts as an agent's evidence *in situ*, and it puts forward a story that explains why things turn out as they do in this respect.

In outline, at least, we may think of the theory of evidential information as springing from the answers to three basic questions about evidence:

1. Do factors which generate evidence have propositional content?
2. Do those factors supervene on the mental?
3. Is evidence always compatible with truth?

Yes-no answers make for an eight-fold space of theoretical possibilities (see Table 4.1).

There are well-known motivations for at least six positions in this space. Theories on line one say that factors which make for evidence manifest propositional content, that those factors cannot be changed without a shift of mental state, and necessarily factors which generate evidential information dovetail with reality. The resulting approach to evidence is favoured by those who say evidence is pinned down by an agent's knowledge, for instance, a signature claim of knowledge-first epistemology.²

Theories on line two say that factors which make for evidence manifest propositional content, that those factors cannot be changed without a change of mental state, and that it is possible for such factors not to dovetail with reality. This sort of approach is favoured by those who say evidence is fixed by an agent's background beliefs (for example). Views of this sort were often baked into twentieth-century epistemology. They are precisely the sort of thing that knowledge-first epistemology is meant to push against.

Theories on line three say that factors which make for evidence manifest propositional content, that such factors can be changed without a change of mental state, and that perforce they dovetail with reality. This sort of approach is favoured by those who say that evidence is fixed by real-world facts, facts fully external to an agent.

Table 4.1 Options for a theory of evidential information

	Manifest propositional content?	Supervene on the mental?	Dovetail with reality?	Evidence
1	Yes	Yes	Yes	Knowledge
2	Yes	Yes	No	Belief
3	Yes	No	Yes	The facts
4	Yes	No	No	?????
5	No	Yes	Yes	Veridical perception
6	No	Yes	No	Conscious experience
7	No	No	Yes	Material objects
8	No	No	No	?????

Theories on line four say that factors which make for evidence manifest propositional content, that they can be changed without a change of mental state, but that factors which make for evidence needn't dovetail with reality. I do not think anyone puts forward a theory of evidence like this, and it is clear why that would be so. After all, any such theory would say that truth-theoretic factors which potentially conflict with reality help fix an agent's evidence, yet it would also say that those factors fail strongly to correlate with an agent's mental states. It is unclear how any such picture could form into a story about the nature of evidence.

Theories on line five say that factors which make for evidence fail to manifest propositional content, that they cannot be changed without a change of mental state, and that perforce such factors dovetail with reality. This is the sort of view favoured by those who say that evidence is fixed by object-seeing. Non-content-based versions of disjunctivism about perception spell out the evidence – in the so-called 'good case', at least – exactly in line with the view.

Theories on line six say that factors which make for evidence fail to manifest propositional content, that they cannot be changed without a change of mental state, but that such factors needn't dovetail with reality. This sort of theory is favoured by those who see evidence as fixed by non-world-involving consciousness – sometimes known as 'raw feels'. The obscurity of the resulting view does not mean it lacks motivation. After all, despite that obscurity, the view crops up repeatedly in the work of students, and shines through a good deal of mid-twentieth-century discussion of perceptual justification – all of which took place before the theory of perceptual consciousness took on a truth-conditional turn – and so on.

Theories on line seven say that factors which make for evidence fail to manifest propositional content, that they can be changed without a change of mental state, and that perforce such factors dovetail with reality. This sort of view sees evidence as fixed by external-world objects like tables, rocks, knives or guns. It is the object-theoretic analogue of views on line three of Table 4.1. Whereas the latter say that evidence is fixed by external-world facts, views on line seven swap objects for facts in their approach to evidence.

Theories on line eight say that factors which make for evidence fail to manifest propositional content at all, that they can be changed without a change of mental state, and that such factors needn't dovetail with reality. On this sort of view non-content-theoretic, non-mental factors which might somehow conflict with reality generate an agent's evidence. It is unclear how a view like this could be true, but notionally the position exists in our conceptual space.

5. Evidential support

Not only does full-dress epistemic theory rely on an approach to epistemic targets, as well as an approach to evidential information, it also relies on a theory of evidential support. This is the bit of theory which details how an agent's total evidential information supports, undermines, or is neutral concerning the truth of pieces of information. And here too the relevant story can be pinned down – in outline, at least

– by answers to three basic questions. This time the questions concern evidential support:

1. Can it be accessed *a priori*?
2. Is the support an on-off matter or does it come in degrees?
3. Is it objective or essentially indexed to an agent's point of view?

These questions generate another eight-fold range of possibilities:

There are well-known motivations for at least six of the possibilities shown in Table 4.2. Theories on line A say that evidential support is *a priori*, that it is coarse-grained in nature, and that it is independent of an agent's point of view. This sort of perspective on evidential support is at work, for instance, when evidential information is said to support a given claim by logically entailing it. Views in the category are often presupposed in classic literature on evidence and explanation.

Theories on line B say that evidential support is *a priori*, that it is coarse-grained in nature, and that it is dependent upon an agent's point of view. This sort of story is at work when evidential information is said to support a given claim for an agent because and in virtue of her believing the claim, say conditionally on the evidential information being true. On this sort of story, what it is for a piece of information to support some claim is itself an indexed matter – in particular it is indexed to an agent believing the supported claim conditionally on the information which supports it. So-called 'belief revision' literature uses an approach like this to evidential support.

Theories on line C say that evidential support is *a priori*, that it is a fine-grained affair, and that it is independent of an agent's point of view. This sort of view is at work, for instance, when evidential information is said to support a given claim by raising its logical probability conditional on the information – in other words, when a piece of information supports a given claim because the latter has a high logical probability conditional on the former being true. Any approach to confirmation which used logical probability would use an approach to evidential support of this sort.

Theories on line D say that evidential support is *a priori*, that it is fine-grained in nature, and that it is dependent solely upon an agent's point of view. This sort of view is at work, for instance, when evidential information is said to support a given claim

Table 4.2 Options for a theory of evidential support

	<i>A priori</i> v <i>A posteriori</i>	On-off v graded	Objective v subjective	E.g.
A	<i>A priori</i>	On-off	Objective	Logical entailment
B	<i>A priori</i>	On-off	Subjective	Conditional belief
C	<i>A priori</i>	Graded	Objective	Logical probability
D	<i>A priori</i>	Graded	Subjective	Conditional credence
E	<i>A posteriori</i>	On-off	Objective	Nomic necessity
F	<i>A posteriori</i>	On-off	Subjective	??
G	<i>A posteriori</i>	Graded	Objective	Chance
H	<i>A posteriori</i>	Graded	Subjective	??

entirely because the agent has high conditional credence for the claim on the information. On this sort of view, what it is for a piece of information to support a given claim is for an agent to lend high credence to that claim when supposing that the evidential information is true.³ Subjective Bayesians use this sort of approach in their work.

Theories on line E say that evidential support is *a posteriori*, that it is coarse-grained in nature, and that such support is independent of an agent's point of view. This sort of view is at work when evidential information is said to support a given claim, for instance, because the evidential information lawfully necessitates the claim. What it is for a piece of information to support a given claim on this view is for the truth of the information to be nomically sufficient for the claim to hold.

Theories on line F say that evidential support is *a posteriori*, that it is coarse-grained in nature, and that it is dependent upon an agent's point of view. This sort of view is at work, for instance, when evidential information is said to support a given claim for an agent when her belief in that claim conditional on the information turns out to be nomically necessary. On this sort of view, what it is for a piece of information to support a given claim for a given agent is for belief in that claim by the agent conditional on the evidential information to be lawfully true. It is easy to see how certain motivations could lead to a view of this sort: simple-minded externalism about evidential support, together with a commitment to subjectivism about coarse-grained evidence, lead to line-F views of evidential support.

Theories on line G say that such support is *a posteriori*, that it is fine-grained in nature, and that it is independent of an agent's point of view. This is the sort of view at work when evidential information is said to support a given claim by virtue of the objective chance of that claim conditional on that information. Of course the position is quite similar to those found at line C. The only difference between them is that C-views insist, while G-views deny, that evidential support be *a posteriori*.

And the last line of Table 4.2 says that evidential support is *a posteriori*, that it is fine-grained in its nature, and that it is dependent on an agent's point of view. This sort of view says that evidential information supports a given claim for an agent to exactly the extent that her credence for the claim, conditional on the information, matches the objective conditional chance of the claim on the information. I don't know of anyone who puts forward such a view but it is easy to locate its philosophical motivations. The desire for fine-grained support relations grounded in something objective, for instance – but also sensitive to an agent's perspective – leads naturally to a view of evidential support characterized by line H.

6. Conversion theory

Not only does full-dress epistemic theory rely on an approach to its targets, an approach to evidential information, and an approach to evidential support, it also relies on a conversion theory. This theory will pull everything together we've discussed to this

point and use it to generate prescriptions for the configuration of an agent's epistemic targets.

Suppose that belief is the basic focus of epistemic evaluation, that evidential information is whatever is believed by an agent, and that evidential support is logical entailment. Then it would make sense to insist that an agent's set of reasonable beliefs is closed by logical implication, and, for this reason, that whenever a rational agent believes things which entail a not-believed thing, then and for that reason, the agent should *come to believe* the not-believed thing. This is precisely what you find in various models of rational belief revision. They presuppose a conversion theory on which epistemic prescription is pinned down by the upshot of a coarse-grained theory of targets, a belief-based theory of evidential information, and an entailment-based theory of evidential support.

Not everyone accepts an approach to conversion theory like this. And the main divide in the area to be emphasized here is precisely the split between those who accept, and those who reject, the idea that epistemic prescription is fixed *solely* by targets, evidence and evidential support. Orthodox conversion theory maintains that epistemic prescription is fixed by these elements of epistemic theory. Unorthodox views insist that more is needed to say how an agent should configure her epistemic targets. Two sections hence we'll see that there are good reasons to endorse the unorthodox approach. Before getting to that we should spell out how the elements of the theory of rationality fit together.

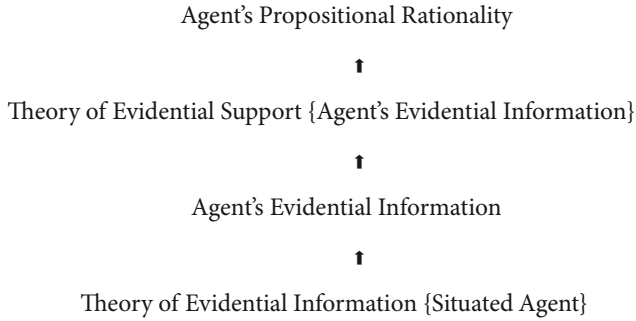
7. Connecting the elements

Drawing everything together, then, full-dress epistemic theory has four major moving parts: an approach to epistemic targets, an approach to evidential information, an approach to evidential support, and a conversion theory. The parts fit together in two steps.

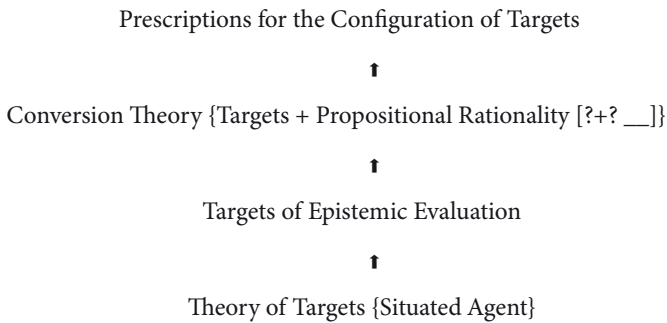
In what I call 'the evidential step', the theory of evidential information pins down which pieces of information belong to an agent's total evidence. The theory of evidential support then kicks in, detailing the extent to which an agent's total evidence supports particular pieces of information. The upshot of this step is an approach to what is known as *propositional* rationality. To a rough first approximation, this is the sort of rationality that is available purely on the basis of the evidence to hand.

In what I call the 'attitude step', the theory of epistemic targets tells us which aspects of an agent's psychology are subject to epistemic evaluation. This is then put with the upshot of the evidential step and the result is fed into a conversion theory. That theory details exactly how an agent should configure her epistemic targets. The upshot of the attitude step is an approach to what is known as *doxastic* rationality. To a rough first approximation, this is the sort of rationality that results from movements-of-mind being driven, causally, by evidentially relevant features of an agent's mental states.⁴

We may picture the evidential step as follows:



And we may picture the doxastic step as follows:



This is how the major moving parts of epistemic theory fit together to generate a detailed story about evidential rationality, and a detailed story about how agents should configure their epistemic targets. In the next section we'll see how knowledge-first epistemology fits into the picture.

8. The epistemic role of knowledge-first epistemology

Consider knowledge-first epistemology and the theory of epistemic rationality. When fitting them together a pair of questions press straight away:

1. Where does knowledge-first epistemology fit into the theory of rationality?

and

2. How does it contribute there?

At least part of the answer to the first question turns on the idea that knowledge is a means of *evidence possession*.⁵ The familiar idea is that an item of knowledge is part of a knowing subject's evidence because it is an item of knowledge. This entails that whatever is known is part of an agent's total evidence:

(1) (K TE).

Through this idea, knowledge-first epistemology contributes to the theory of rationality, and it does so in two ways: it helps to pin down which pieces of information are part of an agent's total evidence, and it helps to clarify why those pieces of information qualify to be part of an agent's total evidence.

On the assumption that there is such a thing as rational credence, moreover – an assumption we'll make in what follows – principle (1) ensures that knowledge plays a serious role in the fixation of rational credence. This will be so because items of knowledge feed into the evidential support function via their role as part of the total evidence; and the upshot of their doing so will itself then feed into the conversion theory. To understand the full role of knowledge-first epistemology in the theory of rationality, therefore, we must consider these latter elements of theory. And here it helps to do things in reverse order.

To begin, the most popular idea about the conversion of propositional into doxastic rationality is so simple that it's often taken for granted in the literature. The thought is that strength of credence should match strength of evidential support on the total evidence:

(2) $cr(\Phi) \cong ES(\Phi, TE)$.

If that is right – if how strongly one should believe something, so to say, is identical to how strongly that thing is supported by one's total evidence – then credal-worthiness can be read off evidential support on total evidence. Knowledge-first epistemology's role in epistemic theory thus turns at least in part on how epistemic support functions work. After all, (1) and (2) jointly ensure that knowledge helps to pin down rational credence by feeding known information into an evidential support function. Any such function will be of this form:

$ES(H, TE) = x$.

The function will take pairs of things as input and yields a unit real number as output. The input pair will itself be built from a hypothesis to be supported by total evidence together with an evidential story fixed by that evidence. The unit real number will represent how strongly the evidence-based story supports the input hypothesis.

Epistemic support functions are a part of epistemology. But they are strongly analogous to similar functions in logic. One thing logic does, after all, is study arguments, and arguments are basically ordered pairs (C, PS) , where C is a conclusion and PS is a premise-based story. A logical support function is thus a function of this form

$LS(C, PS) = x$.

Here we have something which takes arguments as input and yields numbers as output. This time the output represents how well an argument's premise-based story logically supports its conclusion. Support functions in logic are strongly analogous to support functions in epistemology.

There is disagreement about how to make sense of support functions in logic, and a major reason for it has to do with the nature of *validity*. On a general approach, though – to be assumed here and dubbed ‘the risk-based approach’ – validity consists in some kind of guarantee, forged by an argument’s premises, that its conclusion has a minimal risk of being false. The guarantee can be spelled out in modal, alethic, or probabilistic terms – or some combination of them – but the common thread in the work is that validity consists in premises minimizing risk of error.

On the risk-based approach to validity, of course, it is trivial that premises of an argument guarantee a conclusion when that conclusion is one of the premises. This is why arguments automatically get top marks for logical support whenever their premises contain their conclusions. After all, the risk of error in that case is manifestly minimal. Hence we arrive at a fact about logical support functions on present assumptions:

$$(*) \quad (C \in PS) \quad LS(C, PS) = 1.$$

When an argument’s conclusion is part of its premise-based story, the argument provides maximal support for its conclusion.

This suggests that evidential support functions work the same way. After all, it is natural to think there is always as much evidential support as there is logical support, since logical support seems a *type* of evidential support. For this reason, epistemic models of rationality in general – and knowledge-first ones in particular – tend to accept the idea that there is always maximal evidential support on the total evidence for each bit of that evidence. In symbols, they tend to accept

$$(3) \quad (\Phi \in TE) \quad [ES(\Phi, TE) = 1].$$

9. The Input Problem

But now we’re in trouble. (1) and (3) jointly ensure that known items enjoy maximal evidential support on total evidence

$$(4) \quad K(\Phi) \quad [ES(\Phi, TE) = 1].$$

Yet (4) and (2) – the idea, recall, that how strongly one should believe a claim is identical to how strongly it is supported by the total evidence – jointly ensure that certainty is required for known propositions

$$(5) \quad K(\Phi) \quad [cr(\Phi) \equiv 1].$$

That is clearly wrong. Rational agents are not obliged to be sure of things they happen to know. Very often the opposite is true: rational agents are forbidden from being sure of claims they know. In a great many cases this will be so because agents know, as we do, that their sources of knowledge are *fallible*.

Knowledge-first epistemology faces the Input Problem. This a theoretical difficulty for any epistemology which endorses the idea that new evidence comes in bits of

information (i.e. alethic chunks offered by evidential sources of information). Since knowledge-first epistemology endorses the idea that items of knowledge are bits of evidence, and since it also endorses the idea that items of knowledge are alethic chunks, knowledge-first epistemology faces the Input Problem in a clear way.

The problem involves three fundamental ideas: items of knowledge are bits of evidence, credal strength should match evidential strength on total evidence, and evidential information enjoys maximal support on total evidence. If knowing something means an item of knowledge is part of the total evidence, however, and credence for something should match its evidential support on that evidence, and items of evidence are always maximally supported by total evidence, then, maximal credence is always called for when something is known. But maximal credence is not always called for when something is known.⁶

10. Reactions to the Input Problem

On present assumptions at least one of three things must be true: either knowledge does not always yield pieces of evidence, pieces of evidence are not always maximally supported by total evidence, or credence for a claim should not always match that claim's evidential support on the total evidence. It is unclear which option should be taken by knowledge-first epistemology. But we'll assume that the view should resist the first option. This amounts to the assumption that at least part of knowledge's role in epistemology is to connect with the theory of evidential information, and that the connection will ensure that items of knowledge are bits of evidence. This is just a fancy way of saying that we'll assume knowledge-first epistemology endorses something like (1).

Given (1) is true, however – and also that one should not be certain of contingent items of knowledge – one or more of two other things must be true as well. Either

- (A) Knowledge-based items of evidence are *not* always maximally supported by the total evidence; or
- (B) Credence for a claim needn't match in strength the evidential support the claim receives from total evidence.

We'll close our discussion by looking at which of these options knowledge-first epistemology should take.

11. Reflexive evidence?

We're assuming that whenever an item is known it is thereby part of the total evidence:

- (1) (K → TE).

If (A) is true as well – if items of knowledge are *not* always maximally supported by the total evidence – then it looks to be possible for an item of knowledge not to be

maximally supported by evidential information even when that information includes the known item. If evidential information includes a known item, of course, evidence entails the known item. And this means the probability of the known item is maximal on the evidence. So why might we fail to accept – in such evidentially propitious circumstances – that evidential support is maximal? Why might we reject the view that evidential support tracks evidential conditional probability in such a case?

One line of thought presses an analogy with question-begging arguments. To see how it works, consider such an argument in schematic form:

(QBA)	P_1 P_2 $\bullet$ $\bullet$ P_n C

	$\therefore C$

The conclusion of this argument is explicitly one of its premises. Hence the argument explicitly begs the question when it comes to the truth-value of its conclusion. In the minimize-risk sense of logical support, though, (QBA)'s premises maximally support its conclusion. But do its premises *establish* that conclusion?

That depends on what we mean by 'establish', of course. (QBA)'s premises *do* tell a story the truth of which guarantees the truth of its conclusion. Since the argument's conclusion is part of its premise-based story, moreover, the guarantee is automatic. Risk-of-error guarantees have proved useful in logic, probability theory, and other areas of concern. In one useful sense of 'establish', therefore, (QBA)'s premises establish its conclusion. In that sense the argument's premises maximally support its conclusion.

Having said that, it is clear that the explicitly question-begging argument cannot underwrite a rational transition-in-thought *from* a place where C is an open question *to* one where C is not an open question. It is not possible rationally to travel in thought, so to say, by appeal to (QBA), from no view of C (for instance) to a position on which C is true. Since the argument's premises explicitly commit to C 's truth, commitment to the argument's premises brings with it commitment to its conclusion. Rational endorsement of C cannot be *generated* by commitment to (QBA)'s premises. Despite those premises establishing the argument's conclusion, in at least one useful sense of 'establish', the premises are incapable of generating rational commitment to its conclusion.

Explicitly question-begging arguments are incapable of generating rational commitment to their conclusions. When the topic is whether C is true, therefore, (QBA) is a hopeless tool for rationally shifting opinion. Yet the signature role of evidence – its *raison d'être*, so to say – is to function as a driving force behind rational shift of opinion. If evidential claims work like premises of an argument, however, the thought here would be that they are incapable of playing this role, at least when it comes to their own truth-value. Hence it might be thought that no claim can be part of

the evidence when it comes to its own truth-value. It might be thought instead that the very notion of 'total evidence' shifts from claim to claim in a way which rules out an explicitly question-begging set-up. On this perspective it is not automatic that knowledge-based items of evidence are maximally supported by the total evidence, for it is not automatic that they are part of the total evidence, at least when it comes to their own truth-value. When it comes to that truth-value pieces of evidence do not count as evidence at all.

On this way of looking at things, total evidence shifts from question to question, from subject-matter to subject-matter; and the shift always guarantees that no claim is part of the evidence when it comes to its own truth-value. When a claim becomes some of the evidence, for this reason, it is not automatically supported by the evidence concerning its truth-value. The nature of evidential support rules out question-begging-ness in this way.

This line of thought yields an escape route from the Input Problem. But it creates new theoretical questions to be answered. Chief among them is this: how confident should one be in a piece of new evidential information? The knowledge-first epistemology before us – the one committed to (1), (2) and (3) – has the good-making feature of answering this question directly. But it also has the bad-making feature of answering it in an obviously-wrong way. It is clear that we should not always be sure of new pieces of evidential information. What is unclear is how sure we should be of those pieces of information. This is the Credence-for-Evidence problem, something faced by any epistemology which sees evidence to involve pieces of information offered by fallible sources of evidence.⁷

12. Propositional and doxastic rationality

The line of thought sketched in the previous section supports point (A) – the view that knowledge-based items of evidence aren't always maximally supported by the total evidence. The line is based on the idea that evidence is meant to push along rational shift of opinion. It might be thought this shows the line is no good, for it might be thought that the line illicitly runs together propositional and doxastic rationality.

Recall the distinction between them. To a rough first approximation, propositional rationality has to do with evidential relations between contents, and doxastic rationality has to do with the exploitation of those relations – perhaps with other things – in the rational formation of opinion. The push-back against the line of the previous section is that it runs these two types of rationality together, hearing (A) as if it covers both sorts of rationality at once. The thought would be that this is a mistake, since (A) only ever covers propositional rationality, not its doxastic cousin.

Once the distinction between propositional and doxastic rationality is on the table it must be admitted that the line of the previous section appears to involve a conflation. If that is right, though, the line fails to put genuine pressure on the popular idea that claims fully support themselves evidentially. If claims fully support themselves evidentially, however, then collections of claims so support their members, and that entails that one always has full evidential support for each piece of evidential

information to hand. Yet the most elegant and simple conversion theory for credence maintains that strength of credence should match strength of evidential support to hand. Once everything in place here is accepted, therefore, it follows that one ought to be sure of each piece of evidential information to hand.

That is obviously wrong. Self-directed maximal evidential support does not automatically convert to an obligation for maximal doxastic commitment. In some obvious sense that begs the question – not dialectically, of course, since there is no debate between two people here. But the notion of soliloquy is enough on its own to get the idea off the ground, to make sense of the fact that it simply begs the question to convert reflexive-but-full evidential support into obligation of maximal doxastic commitment.

13. Contra naïve conversion

Consider another argument for the view that the Input Problem springs from a naïve conversion theory. It rests on the following claim:

The Variation Hypothesis. There should be no single value v such that, for every piece of evidential information to hand, credence mandated for that piece of information is v .

The idea is that, on any plausible theory of evidential information, variation in credence for pieces of evidence should be permitted. After a bit of reflection this seems pretty obvious, since sources for evidential information vary in their *bona fides*.

It follows from the variation hypothesis that either

- (a) The theory of evidential support treats evidential information variably, giving different values in light of the total evidence to distinct pieces of evidence,
- or,
- (b) Conversion theory violates the injunction to proportion credence for evidence to strength of evidential support it is given by total evidence.

Suppose that the theory of evidential support has only to do with propositional rationality. Then it will not treat evidential information variably. In turn this will mean that conversion theory is unorthodox, violating the injunction to proportion credence to strength of total evidence.

Here's another way to sneak up on the issue. Say that an evidential support function $ES(-, -)$ is *self-addressed* if it is defined for situations like this

$$ES(\Phi, \Gamma) = x, \text{ with } \Phi \in \Gamma.$$

Recall that facts of the form

$$ES(\alpha, \beta) = x.$$

codify how strongly α is evidentially supported by β .

Since evidential support has to do with propositional rationality, which in turn has only to do with how contents are supported by contents, it follows that whenever $\Phi \in \Gamma$

$$\text{ES}(\Phi, \Gamma) = 1.$$

Self-addressed evidential support is always maximal. Intuitively put every claim does maximally well on its own truth. But whenever Φ is a piece of evidential information and Γ is the total evidence, Φ is a member of Γ . So in every such case the evidential support for Φ is maximal.

This point does not depend on evidential support manifesting the structure of a probability function, or the structure of a Dempster-Shafer function, or the structure of any other fancy gizmo found in formal epistemology. Formalization and its limits have nothing to do with the problem. The bother has only to do with evidential support being self-addressed and maximal in the self-addressed case. If that occurs then conversion theory will have to be unorthodox, for it is simply not true that certainty is obliged for each piece of evidential information.

14. The Münchhausen constraint

We should have known this already. To see why: suppose you know Φ_1 on the basis of it looking as if Φ_1 , and you know Φ_2 on the basis of it looking as if Φ_2 , etc., all the way down to knowing Φ_{large} on the basis of it looking as if Φ_{large} . Suppose also that you know how it looks in each case by introspection, you know the number of cases is large by reflection on your general situation, and you're good at creating new knowledge by putting pieces of known fact together.

In the event it is both plausible, and knowledge-first epistemology entails, that you have maximal evidential support for the view that your look-states are fully accurate across a large range of cases:

$$\text{ES}(\text{Large-}\#\text{-of-fault-free-looks/your knowledge}) = 1.$$

After all, [looks as if Φ_1 when Φ_1], ..., [looks as if Φ_{large} when Φ_{large}] are all pieces of evidential information, since you come to know each of these when-claims by appeal to two other pieces of knowledge you have.

Should maximal credence be given in the case – or anything remotely like maximal credence – to the view that your look-states have been faultless over a large range of cases? No: evidential support does not make for doxastic prescription in this way. It is impermissible to adopt a view about the *bona fides* of a source of information on the basis of commitments got solely (or even largely) from that source of information. This is both obvious and connected to begging the question.

I call this the *Münchhausen* constraint on the adoption of attitudes, for flouting it would amount – in good Baron Münchhausen style – to pulling yourself out of a pit

solely by grabbing your own ponytail and yanking upward. No such yanking could succeed, and no flagrant bootstrapping of attitudes about the *bona fides* of an evidential source could be rationally based solely (or even largely) on information drawn from that source.

The Münchausen constraint plays a central role in conversion theory. Within knowledge-first epistemology it ensures that conversion of evidential support to doxastic prescription is unorthodox – or at least it is unorthodox when agents know, as we do, that they are fallibly situated in the world. And two points are relevant here. First, the Münchausen constraint cuts against the idea that strength of credence should be identical to strength of evidential support. In the bootstrapping case sketched earlier, for instance, you have maximal evidential rationality for the view that certain look-states are fully accurate. That evidential rationality derives almost exclusively from commitments understood by you to be got solely from the look-states in question. This means the Münchausen constraint precludes propositional rationality from converting into doxastic prescription about the *bona fides* of the looks-states involved: credence in the case should not be equal to the level of support there is on your total evidence.

Second, the Münchausen case also shows that credence should not always be proportional to evidential support on the total evidence, for the constraint entails that evidential rationality should not automatically be exploited at all in certain cases. Whether such rationality converts into doxastic prescription depends on a number of things: on the content of the claim in question, on the basing-aetiology of commitments the contents of which make for the evidential rationality in question, and so on. No scalar function of evidential support will capture when support converts into doxastic prescription. In turn this means that conversion theory will be deeply unorthodox, for rational credence need not even be proportional to support there is from the evidence to hand.

Notes

- 1 Thanks to Cian Dorr, Ant Eagle, Nick Jones, Mark Kaplan, Maja Spener and Timothy Williamson for discussion.
- 2 To save space in this chapter I will refer only to work directly relevant to knowledge-first epistemology or the Input Problem. See Williamson (2000) and Greenough and Pritchard (2009) for discussion of knowledge-based pieces of evidence.
- 3 §4.8 of Sturgeon (2020) questions whether supposition is really at work here (rather than a *sui generis* kind of conditional commitment).
- 4 See §13.4 of Sturgeon (2020).
- 5 See Greenough and Pritchard (2009), Hossack (2007) and Williamson (2000).
- 6 Williamson defends a view on which (1), (2) and (3) are true (Williamson 2000). In Kaplan (2003) and his contribution to Greenough and Pritchard (2009), Kaplan notes that Williamson is thereby committed to (5), despite this principle being explicitly and repeatedly in Williamson (2000). In Greenough and Pritchard (2009) Williamson replies, but the reply conflates the input difficulty with mostly irrelevant problems to do with formalization. The Input Problem has nothing essential to do with

formalization. It derives from a tension between four ideas: claims are pieces of evidence; when a claim is part of the evidence it is maximally supported by the evidence; strength of credence should be equal to strength of support on the evidence; and sources of evidence are (perhaps known to be) fallible. Plausible assumptions render these claims jointly inconsistent. The Input Problem is to reckon which claims to give up. See Sturgeon (n.d.) for discussion.

7 See Sturgeon (n.d.).

References

- Greenough, P. and Pritchard, D. (eds) (2009). *Williamson on Knowledge*. Oxford: Oxford University Press.
- Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Kaplan, M. (2003). 'Who cares what you know? Critical study of Timothy Williamson's *Knowledge and Its Limits*', *Philosophical Quarterly* 53(210): 105–116.
- Sturgeon, S. (2020). *The Rational Mind*. Oxford: Oxford University Press.
- Sturgeon, S. (n.d.). 'Credence for evidence', unpublished manuscript.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

Perceiving X = Consciousness of Perceiving X. Hossack and Brentano on the Identity Thesis

Mark Textor

1. Introduction¹

Sometimes when you see, hear, taste or smell, you are aware of seeing, hearing, tasting or smelling. How should we think about the metaphysics of such conscious perceptions?

We can approach this question by asking which faculties are involved in such a mental event: Is awareness of perceiving the result of the operation of a perceptual and a distinct further faculty?

Yes, there is a special sense, the inner sense, answered, to name a few, Aquinas (1265ff), John Locke (1689), Thomas Reid (1785), Franz Brentano in his *Psychologie des Aristoteles* (1867) and, more recently, David Armstrong. The inner sense ‘operates’ on the exercises of sight etc. and when it does so we consciously perceive something.

The negative answer to the faculty question has also been proposed in the history of philosophy. On one reading, Aristotle gave the negative answer in *De Anima*. There is one faculty that puts us into contact with colours, sounds etc. and our perceptions. The Scottish philosopher William Hamilton (1788–1856) articulated the negative answer forcefully: ‘[c]onsciousness is coextensive with our cognitive faculties, – and this is again convertible with the assertion that consciousness is not a special faculty, but that our special faculties of knowledge are only modifications of consciousness’ (Hamilton 1860: 207). Consciousness is a ‘super faculty’ whose exercise on different objects constitutes perceiving and awareness.

Now questions about the individuation of faculties are intricate. Hence, it is helpful to pitch our question also at the level of the actualizations of a faculty, that is, the mental activities in which it is exercised. So we can ask: Does awareness of perceiving consist in the existence of a distinct perception/judgement of perceiving? Higher-order theories of consciousness answer *yes*. For example, the higher-order perception theory has it that awareness of perceiving consists in a perception being targeted by a distinct perception of it.

Just as there are proponents of the view that there is not an inner sense and five outer senses, there are philosophers who take consciousness of perceiving and perceiving to be the same mental event or process. Hamilton can again serve to

introduce the idea. According to Mill, Hamilton's thesis that there is only one faculty implies that awareness of perceiving and perceiving are the same thing.

[Hamilton] denies that we have one faculty by which we know or feel, and another by which we know that we know, and by which we know that we feel. These are not, according to him, different facts, but the same fact seen under another point of view'
(Mill 1865: 115).

Here we have one version of the thesis that I will consider in this chapter. The fact

that MT immediately knows (is aware of) of MT's perceiving at t

is the same fact as the fact

that MT perceives at t .

Mill was not a fan of this Identity Thesis and for a long time it was disregarded. It certainly sounds bizarre: consciousness of perceiving seems to be a mental state that concerns perceiving. How can they be the same? But bear with me.

Keith Hossack wrote a series of papers and chapters in which he ingeniously revived the Identity view of perceptual consciousness.² He takes himself to revive insights he found in the work of the Scottish philosopher Thomas Reid (1710–1796) and the Austrian philosopher Franz Brentano (1838–1916). I will set Reid aside for the purposes of this chapter and focus on Brentano. Brentano concluded one of his arguments by saying:

This suggests that there is a special connection between the object of inner presentation and the presentation itself, and that both belong to one and the same mental act. We must in fact assume this. The presentation of the sound and the presentation of the presentation of the sound form a single mental phenomenon: it is only by considering it in its relation to two different objects, one of which is a physical phenomenon and the other a mental phenomenon, that we divide it conceptually into two presentations.

PES: 126–127 [I: 179]

Passages like this can be read in several ways.³ One reading has it that hearing a sound is numerically distinct from consciousness of hearing the sound, but there is a special connection between them: they are 'fused', the hearing 'contributes to the being' of the consciousness of the hearing.⁴ Hossack's reading is different: there are not two mental events that are related by a special relation, but only one:

e_1 = presentation of x = presentation of presentation of x .

(Hossack 2006: 54)

Four years earlier, Hossack (2002: 173–174) built on a similar thesis an explication of the notion of a conscious mental state:

The state x is a conscious state if, and only if, y 's being in $x = y$'s knowledge of being in x .

For example, *being in pain* is supposed to be a conscious state: my being in pain today is supposed to be the same state as my knowledge or awareness of being in pain today.

In this chapter I will assess the plausibility and fruitfulness of different Identity Theses with special attention to Hossack's discussion of Brentano's formulation of the thesis. I will start by giving an initial argument for an Identity Thesis. I will then expound Hossack's take on the Identity Thesis and his criticism of Brentano. I will try to respond to this criticism on Brentano's behalf. My response to Hossack's criticism will make use of one of Hossack's other contributions to philosophy. Hossack (2000: 413) argued that there are no complex things, only simples (atoms) and ways of referring to the atoms jointly. Whether there are only atoms is of no concern for the purposes of this chapter. But there seems to be joint or plural reference – one sign representing many things. I will argue that plural reference is the key to understanding Brentano's Identity Thesis. The defence of the Identity Thesis that I will propose is therefore a testimony to the fruitfulness of Hossack's thought on plural reference.

2. An argument for the Identity Thesis

In Brentano one can find more than one reason for the Identity Thesis. Here I will focus on one reason that resurfaces in Hossack's work. The argument starts from the epistemic premise that consciousness of one's present perception is immediate and infallible knowledge. Brentano went on to argue as follows:

One can now show that such an immediate knowledge of a fact is only thinkable if the knower and the known stand in the relation of identity. While the known is not absolutely necessary, but only known as a mere fact, it is obvious that it has to persist as long as the evident perceiving is directed on it. For otherwise there would be the contradiction of an evident and false judgement. Hence, what we come to know as a mere fact is known as relatively necessary. This would be unthinkable if the known were not identical with the knower because, of the two independent objects, one could cease to be and the other persist unaltered without contradiction.

RP: 228; my translation

If my consciousness of perceiving is infallible knowledge, I cannot be conscious of perceiving and my consciousness of perceiving is illusory. Hence, necessarily: if I am conscious of perceiving, I perceive. Brentano expresses this in the quote above saying that my perceiving is *necessary relative to* my awareness of perceiving. But there are no *brute* necessary connections between distinct contingent existences like mental events. The modal fact that every possibility in which I am conscious of seeing is one in which I see is striking and cries out for an explanation. What is it about the two events that they travel together through logical space? An answer such as 'Consciousness of seeing depends on seeing' is empty if ontological dependence is understood modally (one

thing *cannot* exist without the other). The answer ‘Consciousness of seeing contains seeing’ is at best incomplete: we need a non-modal notion of parthood for mental events and states and to my knowledge no such notion has been provided. To anticipate: I will introduce a notion of parthood later, but according to it the parts we are concerned with here are ‘conceptual’ parts.

In contrast to the two answers considered, the Identity Thesis provides a straightforward explanation of the necessary connection. Consciousness of seeing and seeing travel through logical space together because they are the same thing. Hence, either awareness of perceiving and perceiving are not distinct existences or consciousness of perception is not immediate and infallible. Brentano and Hossack accept the first option, Armstrong (1968: 107) the second option. Armstrong helps us to see why one wants to resist the Identity Thesis. He thinks of consciousness of your present mental life as a form of scanning a situation:

‘It is clear here that the operation of scanning and the situation scanned must be “distinct existences”. A machine can scan itself only in the same sense that a man can eat himself’

(Armstrong 1968: 107).

If we want to say that consciousness of perceiving is identical with its object, the perceiving, we must give up the idea that in consciousness one scans the contents of an inner world.

3. The Identity Thesis and the metaphysics of events

The argument in section 2 suggests that consciousness of perceiving and perceiving are not distinct existences. If they are not distinct, are they numerically the same ‘thing’? Let’s see whether a positive answer to this question can be defended.

Hossack (2006) articulated the Identity Thesis as a thesis about event identity. He argues that mental acts are events that thinkers undergo and present them with something. Brentano would get off the bus at this point: his psychology is one in which mental acts are not identified with respect to thinkers that perform or undergo them. But let’s set this aside. As soon will become clear this difference is not one that matters for our purposes.

Mental acts are a kind of events. What are events? In his answer to this question Hossack relies on Kim’s (1976: 311f) work on the metaphysics of events: an event is a ‘structured complex’ (ibid.) of a subject, property and a time. Kim (and Hossack) represent the identity conditions of such complexes by means of n-tuples of subjects, properties and times.

Which structured complexes are conscious perceptions? The conscious activity of Hossack’s seeing the sun today, E_1 , is supposed to be the complex that can be represented by the following tuple:

$$E_1 = \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } E_1 \rangle$$

The event of Hossack's seeing the sun today is supposed to contain itself as a constituent. Why? Remember the Identity Thesis:

$$e_1 = \text{presentation of } x = \text{presentation of presentation of } x.$$

If Hossack's perceiving (presentation of x) and Hossack's consciousness of perceiving (presentation of presentation of x) are identical, Hossack's seeing the sun is (also) consciousness of itself and hence represents itself. Given that in the Hossack–Kim scheme events are identified with n -tuples, Hossack's seeing the sun today is identified with an n -tuple of a subject (Hossack), the time of perception and the objects perceived. Since the seeing is also directed on itself, the tuple that represents it contains itself. If the tuples are sets, the set that is or models Hossack's conscious perception contains itself as member.

This is a peculiar result. Consider an intuitive understanding of set.⁵ There are some things that you 'lasso' in. When you have lassoed these things, you have formed a set. But the set is not among the things you put a lasso around. The view that there are sets that contain themselves is not inconsistent, but if we introduce such notions as set and member, sets seem to be things that can't contain themselves. Standard set-theory codifies this intuition in the form of the Foundation Axiom that rules out sets that contain themselves.

Is the fact that perceptions are (modelled by) non-wellfounded sets a problem for Hossack? He points us to non-well-founded set-theory. In this theory graphs are used to picture sets.⁶ This way of thinking about sets does not give rise to the intuition codified in the Foundation Axiom. This is a respectable theory that has uses in computer science. There is therefore 'a reasonable hope' that facts can be constituents of themselves without problem.⁷

However, the modelling of the self-directedness of conscious perception as self-containment in the metaphysics of events has given some pause. For the event of Keith's consciously perceiving the sun at 12.35 pm, on 25 September 2020 turns out to have infinitely many constituents. The set 'unfolds' (Aczel) as follows:

$$\begin{aligned} E_1 &= \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } E_1 \rangle \\ &= \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } \langle \text{Seeing, Hossack, the sun, 12.35,} \\ &\quad \text{25.9.2020, } E_1 \rangle \rangle \\ &= \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } \langle \text{Seeing, Hossack, the sun, } \langle \text{Seeing,} \\ &\quad \text{Hossack, the sun, } \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } E_1 \rangle \rangle \rangle \\ &= \dots \end{aligned}$$

James van Cleve is worried about this regress:

'Hossack [...] says there is no reason that the constituency relation needs to be well-founded. I think myself that we should think twice about a stratagem for avoiding an infinite "upward" regress of conscious states that works only by begetting an infinite "inward" regress of conscious states'

(Van Cleve 2015: 31).

The founder of non-well-founded set-theory, Peter Aczel, sided with Hossack. Take the simplest example of a non-well-founded set: $A = \{A\}$. The set is a finite object: it has one member. But it can be pictured by an infinitely descending tree of sets.⁸ We can also construct an infinite sequence of expressions referring to it: $A = \{A\} = \{\{A\}\} = \{\{\{A\}\}\}$ and so on. The same goes for E_1 : there are infinitely many designators of it. But because one does not need to master all of them to be conscious of perceiving the sun, no vicious infinite regress of mental acts ensues.

Hossack (2006: 55) himself sees another problem for the Identity Thesis. Events are identical if, and only if, the tuples they are (that represent them) are identical. Now, take again Brentano's thesis:

$E_1 = \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020} \rangle = \langle \text{Seeing, Hossack, the sun, 12.35, 25.9.2020, } E_1 \rangle$

The second tuple seems to have one element more than the first. Hence, if the Identity Thesis is true, the conscious seeing must be identical to either the seeing, the seer, the seen object or the time. But this is obviously not the case.

Hossack (2006: 55) concluded 'that the option of literally identifying the experience with consciousness of the experience is not open to Brentano'. I think we should resist drawing this conclusion. Rather, we should reject the identification with n-tuples. Hossack himself considers an alternative solution that I think is in fact Brentano's. But Hossack finds it wanting. I disagree with Hossack on this point. In the next section I will start making my case by outlining an independent motivation for the alternative I have in mind.

4. Brentano on Cantor: what is logical (conceptual) division?

Brentano expounded his Identity Thesis by saying that one mental phenomenon has several 'conceptual parts'. We can motivate the idea of a conceptual part by looking at a philosophical puzzle related to the origins of set-theory. In one of his later papers Brentano responded to a puzzling question by Georg Cantor, the founder of set-theory:

33. [...] Cantor once took 2 apples to a meeting of mathematicians and asked how many things he held in his hand. – Not two, he thought, but infinitely many. For the pair of apples were a thing and made with the two separate apples three things. And this trio was a fourth thing and so on.

34. Few will fail to realise that Cantor's claim is a paradox. Mere conceivings [*Blosse Fassungen*] of the same real occurrence under different concepts create no new reality. The pair is not added to the two separate apples as a new reality [...]

41. Logically divided is what is conceived in different concepts.

VD: 379–380; my translation. See also AN: 366

If Cantor has two apples in his hand, he has also one pair of apples in his hand. If the pair is different from the two apples, we have the pair, the two apples and the pairs of the pair and the two apples, the pair of one of the apples and the pair of the pair of the

other apple. And now we can continue counting pairs of pairs. . . . Hence, you seem not to have exactly two, but infinitely many things that are composed of apples in your hand. But how could you hold infinitely many objects in your hand? Are infinitely many objects not rather heavy? Should we conclude that you are infinitely powerful?

Brentano argues that the paradox can be resolved by the theory of conceptual divisions. There are exactly two apples. No infinite power is required to hold them in one hand. But there are infinitely many different ways to conceive the apples. For instance, the concepts [these apples] and [that pair of apples] are different. The one is a plural concept; the other is a singular concept. They both subsume the apples in your hand. For example, one will want to say:

That pair of apples in your hand is nothing but these two apples in your hand.

The difference between *two* apples and *one* pair of apples is a conceptual difference. There are not particular apples and pairs of apples or sets of apples in the world. There are only different ways of conceiving of several apples. It is worth noting that Boolos (1984: 449) motivates similarly the ontological innocence of plural quantification. Speaking of tigers is not referring singularly to a collection, but plurally to the tigers. Difference in mode of reference, sameness in what is referred to.

5. One mental act, different conceptual parts

Brentano's resolution of Cantor's paradox gives us an initial reason to think that some distinctions which seem to be grounded in different properties of different objects are in fact only conceptual distinctions, that is, concern different concepts under which the same objects fall. The basket does not contain two apples and a pair of apples in addition. There are only two apples in the basket to which we apply concepts of collections and plural concepts.

Brentano applied the same strategy in his philosophy of consciousness. According to Brentano, the distinction between perceiving and conscious perception is a conceptual and not a real distinction. In *Psychology* Brentano expounded his view as follows:

The presentation of the tone and the presentation of the presentation of the tone form one single mental phenomenon, *it is only by considering it in its relation to two different objects*, one of which is a physical phenomenon and the other a mental phenomenon, that *we divide it conceptually* into two presentations.

PES: 97–98 [I: 176–179]; my emphasis, in part my translation

His re-statement in a later manuscript is clearer:

Consciousness of seeing and consciousness of something coloured belong to the same act. It is one act that we divide only conceptually by, on the one hand, conceiving of it as having the coloured thing as object and, on the other hand, as having the colour seer as its object.

RP: 191; my translation

Just as there are two apples that may be brought under different concepts, there is one mental phenomenon that can be brought under different concepts. A conceptual division into two presentations does not 'divide' one thing into two like a cake into two slices. There are two concepts of one thing.

How can, for example, the concepts [seeing blue] and [awareness of seeing blue] apply to the same mental phenomenon? Brentano is famous for arguing that intentionality is the mark of the mental. Consider how he phrased the intentionality mark in a manuscript after 1908:

All thinking is, in a certain manner, directed on something as object, *maybe also at once on several objects in a different way.*

VO: 341; my translation and emphasis

One mental phenomenon may have several intentional objects. Indeed any mental phenomenon, Brentano says, always has many objects:

We have in any unitary mental activity always a plurality of relations and a plurality of objects.

PES: 138 [II: 138]

The basic fact is that intentionality is plural. There is one mental phenomenon that has no parts that are not themselves mental phenomena. Just as the plural demonstrative 'these' is directed on several things at the same time, a conscious perception is supposed to have several objects, among them itself. In *Psychology* Brentano gave several independent arguments for this conclusion. For reasons of space, I cannot rehearse them all here. I will only briefly outline one argument that bears directly on our discussion.

Sometimes things strike us as different: the colour I see and the sound I hear strike me as different. I am further in the position to make a non-inferential judgement that registers their difference. Brentano and his students argue convincingly that my awareness of sound and colour as being different cannot be a function of a singular presentation of a sound and a singular presentation of a colour. In order to be aware of the difference between sound and colour, you need to be aware of both:

Only if sound and colour are presented jointly, in one and the same reality, is it conceivable that they can be compared with one another.

PES: 123 [I: 227]

The basic notion we need is 'x presents some things jointly' (*gemeinsam vorstellen*).⁹ There is one thing – the mental act – that presents several things. The mental act cannot be a multiplicity of singular presentations taken together. A multiplicity of singular presentations of A, B, C – whether these are ontologically interdependent or 'loose and separate' – does not amount to presenting A, B, C jointly.

Plural reference is the key to understanding Brentano's view of consciousness. If the same mental act has several different intentional objects, it can be brought under

different concepts as Brentano proposed. 'Hossack's seeing the sun at 12.35 pm on 25 September 2020' and 'Hossack's consciousness of seeing the sun at 12.35 pm on 25 September 2020' are both *partial* descriptions of one and the same act; neither description identifies all objects of consciousness. But each description describes the same act in relation to one of its objects uniquely. Compare: if John is the brother of Tom and the father of Clare, there are two partial, but unique, descriptions of John.

Brentano's student Stumpf described the process by which we form such descriptions as 'abstraction': 'Mental states form such a unity that it is an abstraction when we speak of parts. But the abstraction is no fiction. It is grounded in how things are' (Stumpf 1906/7: 236). Abstracting something is 'to present it to the mind apart from the other properties that usually go along with it in nature' (Bain 1864 quoted in the OED 1989). And that is what is happening in the case of consciousness.

Let us apply the idea of partial descriptions to respond to Hossack's argument against Brentano. Compare a similar case in which one event is related to several objects and can therefore be identified in several partial ways uniquely:¹⁰

John's throwing a stone at t = John's killing the birds Coco and Skittles together at t .
 Hence, John's throwing a stone at t = John's killing Coco (among others) at t .
 Hence, John's throwing a stone at t = John's killing Skittles (among others) at t .
 Hence, John's killing Coco (among others) at t = John's killing Skittles (among others) at t .

Similarly, Hossack's seeing the sun may be the same event as Hossack's consciously seeing the sun. The descriptions 'Hossack's seeing the sun' and 'Hossack's consciously seeing the sun' denote the same event, but one uncovers the intentional objects of the seeing more fully.

6. What's wrong with conceptual parts?

Hossack considered Brentano's conceptual parts view briefly only to set them aside:

Perhaps what Brentano is saying here is that there is only one mental act, which can be conceived of in two different ways; the same mental act falls under the two different descriptions 'presentation of x ' and 'presentation of presentation of x '. But this proposal faces the same difficulty as before; it may explain the fact of actual co-occurrence, but it does nothing to account for the necessity of co-occurrence. Why is that, necessarily, any mental act that falls under the one description also falls under the other description?

Hossack 2006: 56

Hossack demands that a metaphysics of consciousness explains why consciousness is infallible. The conceptual parts view does not meet this demand. Hence, it is discarded.

Brentano has an answer to Hossack's criticism. The descriptions 'the presentation of a presentation of A ' and 'the presentation of A ' are not the most fundamental

descriptions of a conscious mental event in Brentano's philosophy of mind. Brentano opens the chapter after the one in which he presented the conceptual parts view by asking:

Up to now, it has been shown only that we have presentations of them, and if they are in any way conscious, they must naturally be conscious in this way, since presentations are the foundations of all other mental phenomena. Our present problem is whether mental phenomena are merely objects of presentation or whether they can be in our consciousness in still other ways.

PES: 107 [I: 195]

Brentano answered his own question positively: Yes, of course: every mental act is presented *and* the object of an immediately evident acknowledgement, a non-propositional positing. Hence, every mental act has one more informative and one less informative description. Hossack's seeing the sun satisfies the more informative definite description

Hossack's immediately evident acknowledgement of seeing the sun at *t*

as well as the less informative one:

Hossack's seeing the sun at *t*.

Anything that falls under the more informative definite description must also fall under the less informative one because immediate evidence guarantees truth. So the satisfaction of 'Hossack's immediately evident acknowledgement of seeing the sun at *t*' necessitates the satisfaction of 'Hossack's seeing the sun at *t*'. Here it is the epistemic character of one description that does the explanatory work that Hossack demands.

This point sheds light on the shape of Brentano's thought. According to him, the direction of explanation in the philosophy of mind runs from epistemology to metaphysics and not the other way round. We don't explain the immediate evidence of consciousness by appealing to its metaphysics. Brentano criticized Überweg for trying to give such an explanation:

He [Überweg] believes that the perception of a mental act consists in its being combined with other acts. As a result of this combination the real act becomes a part of a cohesive whole which is formed out of the totality of real acts. It is this assimilation into the whole which constitutes the act's being perceived and known. It is necessarily perceived and known as it really is, since it is assimilated in its reality.

PES: 108 [I: 198]

In essence, Überweg argued that consciousness of, say, hearing is immediate and infallible knowledge because consciousness of hearing is *composed of* or *contains* the activity it is about.

Brentano correctly, I think, pointed out that the outlined explanation is empty. We have no independent account of part-hood or composition that makes it plausible that a mental act can have a different part from the one it actually has or that it cannot lose parts. Even worse: we don't know what it is for a mental act to have a part or be composed out of mental acts in the first place. Finally, it is an open question whether mereological and/or metaphysical relations ground or explain epistemic properties. Let's assume that it is true that consciousness of hearing is a complex that has hearing as a part. How does this mereological fact contribute to the epistemic property of immediate evidence that the whole is supposed to have? I can see no convincing answer to this question forthcoming.

Brentano turns the direction of explanation around: epistemic facts come first and the metaphysics of consciousness needs to be compatible with them:

The truth of inner perception cannot be proved in any way. But it has something more than proof; it is immediately evident. If anyone were to mount a skeptical attack against this ultimate foundation of cognition, he would find no other foundation upon which to erect an edifice of knowledge. Thus, there is no need to justify our confidence in inner perception. *What is clearly needed instead is a theory about the relation between such perception and its object, which is compatible with its immediate evidence. And as we have already said, such a theory is no longer possible, if perception and object are separated into two distinct mental acts, of which the one would be only an effect of the other, say.*

PES: 109 [I: 198–199]

It is non-negotiable that consciousness of hearing is immediate and infallible knowledge of hearing. If consciousness of hearing and hearing were distinct existences, consciousness of hearing would not be immediate and infallible knowledge. Hence, consciousness of hearing and hearing are not distinct existences. Here epistemology drives metaphysics.

Brentano's negative conclusion in the passage above is of independent interest. According to the causal theory of mind, the causal profile of the mental is distinctive of it. Now, causal relations obtain between 'distinct existences': if x causes y , y might have existed, although x did not. Mental phenomena are distinct existences, they are 'loose and unconnected'. Brentano points us to a limitation of the causal theory: perceptual consciousness cannot be understood in terms of causal relations. In fact, Brentano will go on to extend his account of the metaphysics of the mental to all simultaneous mental acts. Consider what Brentano said about love:

If the presentation and the love were two separate acts, each a thing in its own right, and if one were only by chance the cause of the other, then it would be possible for this cause to be replaced by another, so that we could love something for which we have no presentation. Therefore in every case the love and the presentation of the object loved must belong to the same real unity.

PES: 158–159 [I: 225–226]

The causal theory of mind is a non-starter.

Now there is some way to go from:

Consciousness of perceiving and perceiving are not distinct existences.

to

Consciousness of perceiving and perceiving are conceptual parts of a real unity.

But Brentano has further arguments that bear on this question. In this chapter my goal was merely to protect Brentano's proposal from an unmotivated explanatory demand. Both Brentano's and Hossack's Identity Thesis are promising contenders. Which of them is the right one remains to be seen.¹¹

Notes

- 1 I have discussed Brentano, consciousness, acquaintance, the soul etc. with Keith since 2002. These discussions have informed my work on these topics. I rarely agree with Keith, but his thoughts on the matter have shaped and changed my understanding of the problems under discussion.
- 2 See Caston (2002) and Kriegel (2003).
- 3 See Kriegel (2018).
- 4 Textor (2006) defends this reading of Brentano. As readers of this chapter will be able to tell, I am no longer entirely convinced by Textor (2006).
- 5 Here and below, I follow Boolos (1971): 220.
- 6 See Aczel (1988): 3ff.
- 7 See Hossack (2006): 55 and (2008): 184.
- 8 See Aczel (1988): 6–7.
- 9 See James (1895) and Russell (1913): 8.
- 10 See also Mackie (1997).
- 11 I presented a predecessor of this chapter at the workshop 'Reviving Rationalism. A Celebration of the Work of Keith Hossack'.

References

- Aczel, P. (1988). *Non-Well Founded Sets*. CSLI Lecture Notes.
- Aquinas, T. (1265ff). *Summa Theologica*. <http://dhspriory.org/thomas/summa/>
- Armstrong, D. (1968). *A Materialist Theory of Mind*. Revised Edition, London: Routledge 1993.
- Armstrong, D. (1981). *The Causal Theory of Mind*. Ithaca, NY: Cornell University Press.
- Boolos, G. (1971). 'The iterative conception of set'. *The Journal of Philosophy* 68: 215–231.
- Boolos, G. (1984). 'To be is to be a value of a variable (or to be a value of some variables)'. *The Journal of Philosophy* 81: 430–449.
- Brentano, F.

- PES = *Psychology from an Empirical Standpoint*, trans. by Antos C. Rancurello, D.B. Terrell and Linda L. McAlister, 2nd edn, London: Taylor & Francis e-Library (2009). Page references in brackets to *Psychologie vom Empirischen Standpunkt*. Second edition (1924) in two volumes, ed. by O. Kraus, Reprint Hamburg: Felix Meiner Verlag 1971.
- AN = *Die Abkehr vom Nichtrealen*. ed. by F. Mayer-Hillebrand. Bern: Franke Verlag (1966).
- RP = *Religion und Philosophie: Ihr Verhältnis zueinander und Ihre gemeinsamen Aufgaben*, ed. by F. Mayer-Hillebrand, Bern: Franke Verlag (1954).
- VD = *Vom Denken und vom ens rationis* (1907–1908). In AN.
- VO = *Von den Objecten* (1908). In AN.
- Caston, V. (2002). 'Aristotle on consciousness.' *Mind* 111: 751–815.
- Hamilton, W. (1860). *Lectures on Metaphysics and Logic. Vol I–IV*. Edinburgh/London: William Blackwood and Sons.
- Hossack, K. (2000). 'Plurals and complexes.' *The British Journal for the Philosophy of Science* 51: 411–443.
- Hossack, K. (2002). 'Self-knowledge and consciousness.' *Proceedings of the Aristotelian Society* 102: 163–181.
- Hossack, K. (2006). 'Brentano and Reid on consciousness.' In M. Textor (ed.), *The Austrian Contribution to Analytic Philosophy*. London: Routledge, pp. 36–63.
- Hossack, K. (2008). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Kim, J. (1976). 'Events as property exemplifications.' In M. Brand and D. Walton (eds), *Action Theory*. Dordrecht: Reidel, pp. 310–326.
- James, W. (1895). 'The knowing of things together.' *The Psychological Review* 2: 105–124.
- Kriegel, U. (2003). 'Consciousness as transitive self-consciousness: Two views and an argument.' *Canadian Journal of Philosophy* 33: 103–132.
- Kriegel, U. (2018). 'Brentano's dual-framing theory of consciousness.' *Philosophy and Phenomenological Research* 97: 79–89.
- Locke, J. (1689). *An Essay Concerning Human Understanding*. Ed. by P. H. Nidditch, Oxford: Clarendon Press 1975.
- Mackie, D. (1997). 'The individuation of actions.' *The Philosophical Quarterly*: 38–54.
- Mill, J. St. (1865). *An Examination of Sir William Hamilton's Philosophy*. London: Routledge 1996.
- Oxford English Dictionary (OED) (1989). Definition of 'abstraction'. Oxford English Dictionary 2nd ed., Oxford University Press, <https://www.oed.com/oed2/00000895;jsessionid=4046384CA5DCDA4F3B699D9E4368ECB3>
- Reid, T. (1785). *Essays on the Intellectual Powers of Man*. D. R. Brookes and K. Haakonssen (eds), Philadelphia, PA: Pennsylvania State University Press 2002.
- Russell, B. (1913). *Theory of Knowledge. The 1913 Manuscript*. London: Routledge 1992.
- Stumpf, Carl (1906/7). *Psychologie Vorlesungen Wintersemester 1906/7*. Transcript by W. Koffka.
- Textor, M. (2006). 'Brentano (and some neo-Brentanians) on inner consciousness.' *Dialectica* 60: 411–432.
- Van Cleve, J. (2015). *Problems from Reid*. Oxford: Oxford University Press.

Facts, Knowledge and Knowledge of Facts

Bernhard Weiss

Keith Hossack's book, *The Metaphysics of Knowledge*, is a wonderful presentation of a bold metaphysical system.¹ Hossack welds together insights of recent and modern philosophy to construct his own very individual view on a host of central philosophical concepts: knowledge, facts, universals, particulars, 'concepts, truth, necessity, consciousness, persons and language' (2007: xi). This is, without doubt, a rich offering. Knowledge occupies a pivotal and fundamentally explanatory position; it is seen as a relation between a mind and a fact. Here, with rather less system and organization than Hossack, I attempt to put some pressure on aspects of the encompassing view. We need an account of the ontology of facts in order to understand what Hossack has to say about knowledge. As we'll see, the combination relation plays a crucial role here, but I argue it is impossible for us to comprehend it and so to comprehend the ontology of facts. I then raise some concerns about the strength of Hossack's motivation for treating knowledge as a relation to a fact. Finally, I consider one of Hossack's attempts to recruit knowledge as explanans: his attempt to explain necessity in terms of a priority, raising concerns about his rejection of two counterexamples.

1. Vectors and combination

Hossack introduces the notion of a vector, which is any entities taken in a particular order. Some vectors are entities in the right order to 'make sense'. They make sense precisely when the elements of the vector are combined either in a fact or its corresponding negative fact. So:

$$\text{Sense } \theta =_{\text{df}} (p)(p \bullet \theta \vee p \bullet (\neg, \theta))$$

Here 'p' is a variable ranging over facts, θ is a vector, ' $\neg$ ' is the universal *negation*, and ' $\bullet$ ' is the multigrade combination relation. So we begin with a domain of facts, each of which has components and, what Russell called a relating relation (1903: 50). The relating relation relates the other terms of the fact. So *wisdom* is the relating relation in the fact that Socrates is wise; *love* is the relating relation in the fact that Plato loves Socrates and so on. Universals such as *wisdom*, what Russell in *Principles of Mathematics*

called concepts, can occur either as relating relations or as what he called terms. Particulars such as Socrates can only occur as terms. Russell called these ‘things.’²

The convention that Hossack adopts is that the relating relation is the first element of the vector. So (the fact that Socrates desires wisdom)•(desire, Socrates, wisdom), because desire is the relating relation; but it is not the case that (the fact that Socrates desires wisdom)•(wisdom, Socrates, desire). Given this convention, one can define what it is to be a particular (Russell’s thing), since a particular can never appear in a fact as the relating relation. So:

$$x \text{ is a particular} =_{\text{Df}} (\forall \theta)(\text{nonsense}(x, \theta))$$

The convention could, of course, have been different. Weiss, in contrast to Hossack, chooses to find the relating relation at the end of the vector. So for Weiss the definition of a particular is:

$$x \text{ is a particular} =_{\text{Df}} (\forall \theta)(\text{nonsense}(\theta, x))$$

But what is the convention? The convention is not a convention about vectors, which are just pluralities of entities in order, but a convention about combination. The Hossack convention construes combination as taking the first element of the vector as the relating relation; the Weiss convention construes combination as taking the last element of the vector as the relating relation. In order to understand the multigrade relation of combination, one needs to know which convention is in operation. But, in order to know that, one needs to know what it is for something to appear as the relating relation. So, I conclude that Hossack manages to define being a particular (being such as never to be a relating relation) in terms of a notion of combination which presupposes the notion of a relating relation. This is no metaphysical insight.

The way Hossack presents the theory, it seems that, apart from the first element of the vector, all the others are treated as on a par. So in the fact that Socrates desires wisdom, the elements which don’t appear as relating relations are Socrates and wisdom, and these simply appear as relata taken in a particular order, ensuring that it is Socrates that desires wisdom; rather than that wisdom desires Socrates. I think this is important because Hossack modifies traditional logic to allow for the multigrade relation of combination, but wants a univocal such notion. But this seems impossible, and impossible for reasons that frustrated Russell’s attempts in this direction a full century and more ago. As Russell noted, there are propositions and corresponding facts, which seem to include two verbs.³ These are cases where a proposition seems to include another as component. Russell was exercised by propositional attitude ascription; but the point would apply to logically complex propositions too.

Let’s look at the latter first and adopt Hossack’s preferred mode of framing things in terms of facts. Consider the fact, *p*, that Desdemona does not love Cassio. According to Hossack, $p \bullet (\neg, \text{Love}, \text{Desdemona}, \text{Cassio})$. Here negation appears as the relating relation. But then the problem is that after this first place in the vector the other entities appear simply as ordered entities. And this loses the vital fact that what is negated is something in which *love* is supposed to occur as the relating relation, and as relating

Desdemona to Cassio, in that order. Obviously *love* doesn't occur in any such fact, because Desdemona does not love Cassio. So the temptation is to see it as playing this role in a proposition; it is the proposition which is negated. But this leads to an ontology of propositions which Hossack wants to avoid.⁴ The point is that the combination relation in this instance seems to need to accommodate the fact that the second element is a potential relating relation, claimed to relate the ensuing entities. But then combination is functioning differently here from the way it functions in, say, (the fact that Socrates desires wisdom)•(desire, Socrates, wisdom). To be explicit, in the latter case combination involves taking desire as the relating relation, relating Socrates and wisdom in that order. But in the former case we cannot analogously claim that combination involves taking negation as the relating relation relating *love*, Desdemona, Cassio in that order. Rather we need to say something like combination involves taking negation as the relating relation relating *love* as a relation between Desdemona and Cassio, taken in that order. As Russell pointed out, we cannot treat *love* as an entity on the same level as Desdemona and Cassio; its (potential) role in relating them needs to be preserved.⁵ But then combination is not univocal.

Here's another presentation of the same point. It may well be the case that,

$q \bullet (\text{loves, Cassio, Desdemona})$

$p \bullet (\neg, \text{loves, Desdemona, Cassio})$

In the first q combines *love*, Cassio and Desdemona, taken in that order. In the second p combines *negation*, *love*, Desdemona and Cassio taken in that order. Concentrate on the second place of the combination relation. In the first this is occupied by Cassio, being taken to occupy the first place of the relating relation. In the second it is occupied by *love* being taken as a relation between (here) Desdemona and Cassio; it isn't simply being taken as an entity. So the significance of occupying the second place of the combination relation differs between the first and second cases. The combination relation is not univocal. Let's focus now on the third place of the combination relation, which, as it happens each involve Desdemona (though this is irrelevant to the point). In the first, the significance of Desdemona occupying this place is that she is appropriately combined in the relating relation of the fact, q , namely *love*. In the second, it is that she is combined in a fact p , by the relating relation, *negation*, as the first place in a relation (here *love*). The significance of the third place of the combination relation in either case is different. So combination is not univocal.

The point is that once *love* appears as a mere entity, there is nothing to say *how* it is being construed in relation to the other entities in the vector; but it is being construed in a particular way and the only way to capture this – by reading different significance into the places of the combination relation – demolishes the idea that there is *a* combination relation; there are many.

An alternative would be to shift the burden from combination onto negation. Negation would become a multigrade relation **combining**⁶ the ensuing entities. But then it is unclear to me that negation can be a universal apt to appear in the combination relation. So we might symbolize this as: $\neg(\theta)$, where ' $\neg$ ' is a kind of combination which doesn't lead to fact formation; precisely not. This seems to conjure an ontology of

negated propositions, but doesn't really solve the problem of having the places in the combination relation having different significance in different contexts. And this just for the reason that negation can be iterated. If the first place in θ has a relating significance in $\neg(\theta)$, then this cannot be lost and so the second place will retain this significance in $\neg(\neg, \theta)$. And so on. The effect is that **combination** will mean different things in different contexts; it will not be univocal.

Let me consider other complex facts. Essentially the same problem recurs. Assume the following:

$p \bullet (\text{wisdom}, \text{Socrates})$

$q \bullet (\text{rarity}, \text{wisdom}, \text{beauty})$

So p is the fact that Socrates is wise; q the fact that *wisdom* is rarer than *beauty*. What is apparent is that wisdom appears differently in each of these facts. I think Hossack may want to treat the conjunction as a conjunction of these facts, taken as entities. But this won't work with the disjunction,⁷ since there might be a disjunctive fact even when either disjunct may not be a fact. So, like negation we'll need to consider this as a multigrade relation, perhaps:

$r \bullet (\vee, \text{wisdom}, \text{Socrates}, \text{rarity}, \text{wisdom}, \text{beauty})$

But how to capture, (i) that the second and third; fourth, fifth and sixth elements respectively are to be combined in 'propositions'; and (ii) that *wisdom* appears differently in the second place from how it does in the fifth? If we make this a feature of the combination relation then it ceases to be univocal; if of the universal disjunction, then it ceases to be a universal.⁸

Metaphysically facts are basic. But our comprehension of the ontology of facts accrues through an understanding of vectors and combination. My doubt about whether we have a grasp of combination – because legitimate combination proliferates – is thus a doubt about our comprehension of the domain of facts. Let us bracket this concern and move on to Hossack's claim that knowledge is a relation between a mind and a fact. If he's right about that claim, then, if we're to understand knowledge, we had better, despite my worries, find a way to make sense of facts.

2. Knowledge and facts

For Hossack knowledge is basic; it is the fundamental explanandum in terms of which mental states such as belief are explained and in terms of which truth is explained. Nonetheless, we are presented with an account of the logical form of knowledge ascriptions:

'S knows that A ' =_{df} (x) (p) (x is a mental act $\wedge$ p is a fact $\wedge$ content (x) = that- A $\wedge$ that- A is a mode of presentation of p $\wedge$ S knows of p in virtue of x)

(Hossack 2007: 7)

I have to confess to being unclear about the role of the quotation marks on the left-hand side; and so shall ignore them. I take it thus that we could happily remove the quotation marks and replace ‘ $\stackrel{df}{=}$ ’ by ‘iff’ if we so chose. The RHS presents a bewildering ontological salad: facts, mental acts, contents and modes of presentation. In addition, it helps itself to quantification over facts and over objects; picks out contents by means of that-clauses; and talks about knowledge *of* facts in virtue of being in a mental state. We later discover that the ‘in virtue of’ locution is causal, not constitutive.

Hossack distinguishes knowledge from propositional attitudes, such as belief, with which it is often lumped. He claims there is a fundamental difference between them. The difference emerges out of the factivity of knowledge. Since the true propositional attitudes are not factive they can be taken to be relations to contents. Not so knowledge. One of the ways he makes this point is by means of the knowledge-of relation. A sentence of the form ‘S knows of ...’ is completed by a sentential nominalization. And the knowledge-that claim entails a knowledge-of claim. So ‘S knows that Fred is red’ entails ‘S knows of Fred’s being red’. But there is no corresponding relation between a believes-that claim and a believes-of claim. Indeed, there is no believes-of claim at all.

I don’t find this persuasive. Hossack’s examples of propositional attitude verbs are hopes, beliefs and fears. But these seem to me to be good inferences:

S hopes that Fred is red.

S hopes for Fred’s being red.

S fears that Fred is red.

S fears Fred’s being red.

S believes that Fred is red.

S believes in Fred’s being red.

I don’t know why English alters the preposition with the propositional attitude verb, or does away with it altogether, but am near certain that there is no logical significance to this.

Hossack uses his observation to distance the logic of belief attributions from knowledge attributions; and to argue that the fundamental relation, knowledge-of, is referentially transparent and thus a relation to a fact. If my observation is correct, then the disanalogy between the logic of belief and knowledge attributions is yet to be shown. Also we undercut the support for thinking that knowledge-of is a relation to a fact. For it seems possible to say that S believes in Fred’s being red, though Fred is green.⁹ So the relation spelled out here – between S and the referent of the sentence nominalization – is not a relation to a fact; since Fred is not red. So we have no more reason to think that the relation between S and the sentence nominalization in ‘S knows of Fred’s being red’, is a relation to a fact. If the latter is factive whereas the former is not, this need not be a product of the nature of the relatum – that it is a fact – but of the nature of the relation.

3. The *a priori* and necessity

What Hossack wants to conclude is that knowledge is a relation between a mind and a fact. But surely his account of the logical form of knows-that statements establishes something quite different, namely, that knowledge is a relation between a subject and a content. Of course the 'analysis',¹⁰ quoted above, breaks this relation down into one holding between a subject who is in a mental state whose content is a mode of presentation of a fact. But that doesn't show that knowledge is a relation to a fact. No more does it do so than does the following establish that being a neighbour of is not a relation between people but a relation between people and abodes:

A is B's neighbour iff $(\exists x, y)(x \text{ is A's abode} \wedge y \text{ is B's abode} \wedge x \text{ is near } y)$

Of course, it may well be that the *knowledge-of* relation, rather than the knowledge-that relation, is a relation between a mind (or a subject) and a fact.¹¹ But that's to change the subject. This matters.

To show that it matters let's consider what Hossack has to say about the *a priori* and necessity. His claim is that what is necessary is *a priori*, a Rationalist Thesis:

A fact is necessary if and only if it has an *a priori* mode of presentation.

Prominent counterexamples to the claim are identities, which are claimed to be (metaphysically) necessary but *a posteriori*; and statements involving descriptive names which are claimed to be contingent but *a priori*.

He begins his discussion by citing Leibniz, Kant and Frege as co-travellers, and impressive companions on the road they would be too. But let's register that Leibniz is identifying necessity and *a prioricity* among truths, Kant and Frege among judgements. Hossack is interested in *facts*. This is a little disconcerting. We seek a connection between the categorization of facts and modes of knowing. The resolution comes quite quickly because, in terms of Hossack's conception of cognitive faculties, we can say that 'a fact is *a priori* if it is apt to be known by reason alone'. There is an assumption in this seemingly innocuous statement. The assumption is that it is facts that are known. But this is an assumption that Hossack claims to have provided ample argument for. And so he has; but his argument related to knowledge-of, not knowledge-that. And as I've just pointed out, by his own account, 'S knows that A' spells out a relation between a knower and a *content*.

Let's consider how he deals with the counter-examples. First, necessity without *a prioricity* in some identity statements, such as 'Hesperus is Phosphorus'. Hossack here appeals to Frege's distinction between sense and reference. The fact that Hesperus is Phosphorus is the same fact as the fact that Hesperus is Hesperus; so if the former is necessary and *a priori* so is the latter.¹² The two sentences simply correspond to different modes of presentation of the same fact and, since that fact has an *a priori* mode of presentation, the fact is *a priori*. Though this is what Hossack says, it can't be quite what he means since we don't know what it is for a mode of presentation to be *a priori*. What he must mean is that the fact can be known by reason alone under some mode of

presentation. But facts are not the objects of knowledge-that, so this can only mean that a fact is *a priori* if it can be known-of by reason alone, under some mode of presentation. But since knowledge-of is always independent of mode of presentation, since it is transparent, talk of mode of presentation is here otiose, or, worse, misplaced. But now we begin to lose a grip on the Fregean solution. Insofar as we were interested in modes of presentation (which are contents) then we are interested in knowledge-that, that is, knowledge whose object is a content. And the claim is that knowledge that Hesperus is Phosphorus, is not *a priori*, but is knowledge of a content which, if true, is necessarily true. Hossack's solution is to change the subject; to shift us from concern about knowledge-that and to take us from a concern at the level of sense to one at the level of reference. This isn't a solution; it is to change the subject.

There's some irony in Hossack's deployment of Frege here. Frege distinguished sense from reference in order to use sense to account for cognitive value; and so to feed into accounts of cognitive concepts such as *a prioricity*. Hossack appeals to the distinction but makes the level of sense redundant (as I just mentioned); and then focuses the epistemic concept of *a prioricity* on the level of reference, i.e. in the realm of facts. The trick is, I think, pulled off by failing properly to distinguish knowledge-of from knowledge-that. The philosophical literature is concerned with knowledge-that; but this clearly is ignored in Hossack's discussion when it is properly opened up.

Of course, there's no reason to object to Hossack's relocation of the focus. But then the counterexamples he raises are actually simply beside the point. And, since he is concerned with facts and knowledge-of, he is mistaken to recruit Leibniz, Kant and Frege as fellow companions. These philosophers were concerned with knowledge-that (with truths and with judgements, respectively). But more importantly, it isn't a counterexample that we need him to address; it is a gap. We need to understand how Hossack's notion of *a prioricity* relates to knowledge-that; as yet this has not emerged from obscurity.

Let me now consider the case of examples which appear to show that we may have *a priori* knowledge of contingent truth. Such examples appear plausible because it seems one may fix the reference of a name by using a distinguishing, though contingent, property of an individual. Then the sentence ascribing that property to the individual as picked out by that name is known *a priori* – since it is a stipulation – but is contingent, since it ascribes a property to that individual, which it needn't have had. The example we'll use is Evans's as used by Hossack too. The descriptive name is 'Julius' and the sentence used to stipulate a reference – and thus give a meaning to the name – is: 'If anyone invented the zip then Julius is the inventor of the zip; and if no-one invented the zip then Julius is the number 0'. From this it is claimed that the stipulator, S, knows *a priori* that 'Julius, if anyone, invented the zip'. However since it is possible that Julius might have followed a different career and that Jason had better luck in his project with fasteners, there is a possible situation in which the sentence is false, because in that situation Jason invented the zip and Julius did not. So the sentence 'Julius, if anyone, invented the zip' is contingent, since, though true, is false in the situation just described where Jason, not Julius, invented the zip.

Let's begin by considering what Hossack says about what he calls the Disquotational Argument, which he attributes to Kitcher. The form of the argument is given as follows:

- (1) S knows *a priori* that 'Julius, if anyone, invented the zip' is true.
 - (2) S knows *a priori* that if 'Julius, if anyone, invented the zip' is true then Julius, if anyone, invented the zip.
 - (3) Therefore, S knows *a priori* that Julius, if anyone, invented the zip.
- (See Hossack 2007: 153: I have used Hossack's exact wording,
but changed his numbering of the claims)

This is, I take it, supposed to be a rendering of Kitcher's point. Here's the quotation from Kitcher, given by Hossack:

[S]peakers of English apparently know (*?a priori*) that 'If Shorty exists then Shorty is a spy' is true if and only if, if Shorty exists then Shorty is a spy. So it is hard to see why I can't at least come to know that if Shorty exists then Shorty is a spy, on the basis of the knowledge I obtained through my reference-fixing.

Kitcher 1980: 96; quoted in Hossack 2007: 149

Let me notice straightaway this difference between Hossack's version of the argument and the thought suggested in the quote from Kitcher. Hossack presents a putative argument to show that S, the stipulator, has *a priori* knowledge. But what Kitcher is suggesting is something rather different; he is suggesting that there is a first person mode of reasoning which delivers the requisite (*a priori*) knowledge. So what we really need to be reconstructing is this method of reasoning; that is, we need the first personal version of the argument. Having registered this concern let's push on to uncover Hossack's complaint. It is this.

Hossack begins by noting that the truth predicate ought to be relativized to a language, in this case the language spoken by S. He symbolizes this using Neale's notation for definite descriptions as [the *L*: S speaks *L*]. He then notes that (1) is ambiguous according to whether the definite description is given wide or narrow scope reading:

- (1a) S knows *a priori* that [the *L*: S speaks *L*] ('Julius, if anyone, invented the zip' is true-in-*L*).
- (1b) [the *L*: S speaks *L*] S knows *a priori* that ('Julius, if anyone, invented the zip' is true-in-*L*).

According to Hossack, the narrow scope reading (1a) is true whereas the wide scope reading (1b) is false. I think the reasoning for this is that S knows (*a priori*) that in the language she actually speaks the Julius sentence is true – since she has stipulated this; but in that language, the language she actually speaks, she doesn't know (*a priori*) that the Julius sentence is true; since there are alternative possibilities, consistent with what she knows, in which someone other than Julius – Schmidt or Jason – invented the zip.

Again, we get two readings – wide and narrow scope – for the second premise, as follows:

- (2a) S knows *a priori* that [the *L*: S speaks *L*] (if 'Julius, if anyone, invented the zip' is true-in- *L* then Julius, if anyone, invented the zip).
- (2b) [the *L*: S speaks *L*] (S knows *a priori* that if 'Julius, if anyone invented the zip' is true-in- *L* then Julius, if anyone, invented the zip).

Here the wide scope reading, (2b), is true. It is true because in the language which S speaks, namely *L*, the disquotational schema holds; so this instance holds and S can know this (*a priori*). But the narrow scope reading, (2a), is false. It is false because there is an alternative situation in which Schmidt or Jason invented the zip, and S's stipulation brings it about that, though the sentence 'Julius, if anyone, invented the zip' is true, Julius did not invent the zip (Schmidt or Jason did).

So the complaint is that in order to read the argument as having true premises, the definite description needs to have narrow scope in the first premise and wide scope in the second premise and this leads to an argument with an invalid form. The complaint is subtle and ingenious. Nonetheless its conclusion must be false. For what the argument does is frustrate semantic descent from a stipulation, construed as metalinguistic – to the object language. The stipulation is metalinguistic since it is stipulation that a certain sentence is true. But, when we try to use facts about disquotation to achieve knowledge stated in the object language, Hossack objects that we are equivocating in our scope distinctions. And this means that, though S may make the stipulation, she cannot thereby give herself any entitlement to *use* the term whose meaning she has stipulated. There are two potential ways to resist having to endorse this impossible conclusion. The first is to construe stipulation in a way that sees it as purely an object linguistic phenomenon. The second is to find a way, despite Hossack's argument, of allowing semantic descent. I'll consider the first option first.

Let's first note that there is something implausible about Kitcher's disquotational argument as a construal of the mechanics of stipulation. For, if this were the right way to understand the process, then it would be impossible for a stipulator to be successful, if she failed to have the ability to frame her premises by means of semantic ascent, relevant semantic principles, an understanding of the disquotational schema for the relevant language, and a means of referring to languages. This surely cannot be right. Stipulation is a much more basic feature of language use than any of these others; and so must be possible in their absence. Of course, it may be that the disquotational argument provides some kind of rational reconstruction of the process of stipulation, and I'll return to that suggestion in a moment.

But let me press on by thinking about why it might be that the disquotational construal seems attractive. The question is how best to characterize the act of stipulation. Stipulating seems an act focused on a *sentence* rather than a *content*; and because of this it seems natural to see it as metalinguistic. We don't want to say, for example, that S stipulated that Julius, if anyone, invented the zip, because by means of her stipulation S introduces 'Julius' into her language and, though we might want to report the stipulation, there's no reason for us to be a party to it. So the name 'Julius' might be unavailable to us in reporting what S did. Moreover, it may be, for instance, that we've stipulated a meaning for 'Jason' in precisely the same manner as S has done for 'Julius'. So 'Jason' and 'Julius' are synonyms. So if it's fine for us to say 'S stipulated that Julius, if anyone,

invented the zip' then it would be fine for us to say 'S stipulated that Jason, if anyone, invented the zip'. But the latter would clearly be an inadmissible report of S's stipulation. So there's a natural inclination to report S's act of stipulation metalinguistically as: S stipulated that 'Julius, if anyone, invented the zip' is true. But then we are led to the implausibilities I recently described and, if Hossack is right, to an account in which stipulation fails to sponsor an object linguistic use of a term.

So we need to find a middle way between reporting the content of S's act by means of semantic ascent and reporting it object linguistically, and so being sensitive only to content and not mode of expression. I don't think we have a great mystery here. We simply need to use direct speech: S uttered 'Julius, if anyone, invented the zip' stipulatively. Of course, there is a question about which of the words in that sentence is the subject of the stipulation; but that will be made clear by context and is, equally, a feature of the metalinguistic account. I don't think anything of any moment hangs on this.

Consider the following. U overhears talk involving the name 'Julius' but does not understand the name. So U asks S who Julius is. It is clear that, in doing so, U doesn't understand the name 'Julius' and is asking for an explanation of its meaning. S responds, 'Julius is the inventor of the zip'. Simultaneously, U learns the meaning of 'Julius' and a fact about him, namely, that he invented the zip. The reason for this is that the former suffices for the latter, and thus the latter is *a priori* knowledge.¹³ Nothing changes when the situation is one in which there is no U, and S's statement of 'Julius is the inventor of the zip' is a stipulation laying down the meaning of 'Julius'. S thereby gives 'Julius' a meaning and learns something *a priori* about Julius.

In each case – the explanation of meaning; and the stipulation of meaning – S uses a sentence assertively. In the stipulative case, she has an unearned entitlement to make the assertion. Since the entitlement is unearned, we may well want to call it *a priori*. And, if the knowledge rule for assertion is correct, then the assertion (if correct) expresses a piece of knowledge that S has. Thus S has a piece of *a priori* knowledge. Given that the name 'Julius' is intended as a personal proper name, the metaphysics of people then kicks in to show that this piece of knowledge is contingent.

So, to return to my worry about moving away from Kitcher's first personal construction of the argument, what we need is to find S's passage from the stipulation to the claim. But, if I am right, this is immediate. The stipulation *is* the claim. No inference is necessary; and what inference there is, is the mere stuttering inference.

I suppose that one might contest the claim that the stipulative utterance is a genuine assertion. However, it seems clear that in the case of explaining meaning the utterance is a genuine assertion and I can't see a reason to drive a wedge between the two cases. When S makes her utterance to U, U only learns the meaning of 'Julius' if she takes S to be making an assertion (and a true one at that¹⁴). If U takes S to be making a supposition, merely entertaining a thought, she fails to learn a meaning for the name 'Julius'. So a successful explanation of meaning requires S to be making a kind of assertion; in explaining a meaning, as in assertion, one commits oneself to the truth of one's utterance; likewise in the case of stipulation. Moreover, it is certainly true, in either case, that ensuing uses of the same sentence can be assertions or that valid deductions therefrom are assertions, and are good assertions. S, in the stipulation case, or U, in the

explanation case, can go on to assert this sentence: 'Either no-one invented the zip or Julius invented a fastening device'. Such assertions are warranted *a priori*.

Another tactic would be to distance assertion from the knowledge rule and to claim that in these cases the speaker asserts on the basis of an *a priori* entitlement that doesn't amount to knowledge. I find it hard to see how to apply this tactic to the ensuing practice I just described and difficult to see how it plays out in a practice of inference which will invoke a mixture of claims, some of which are known and some of which are merely *a priori* entitled. At least, I fail to see how to understand all of this, *unless* one gives up on the idea that knowledge is fundamental and uses some other epistemic concept as both fundamental and unifying, perhaps that of warranted assertion. Though, I might be happy with this suggestion, we know that Hossack would not be.

I'll rest my case; but would like to address the question of whether the disquotational account can offer a kind of rational reconstruction of stipulation and Hossack's argument against it. I suppose too, that one might agree with the implausibilities I've mentioned about this construal of stipulation, but think too that, *if one had the requisite concepts*, then one ought to be able to make a stipulation metalinguistically.¹⁵

A crucial part of Hossack's argument here makes use of the idea of an epistemic alternative. He introduces the notion of an epistemic alternative as follows,

w is an epistemic alternative for *S* at *w* iff everything *S* knows at *w* is true at *w*.

As he notes, the notion is useful in arguments for ignorance. Given a case of putative knowledge that, say, *A*, in *w*, then, if there is an epistemic alternative, *w*, for the alleged knower at *w* at which *A* is false then the alleged knower doesn't know *A*.

Let me alter the account slightly. According to Hossack if *w* is an epistemic alternative for *S* at *w* then, if *S* knows *p* at *w*, then *p* is true at *w*.¹⁶ I think it is plausible to enlarge the set of things that need to be true at *w* to include not merely things known by *S* at *w* but anything which follows logically therefrom or is necessarily true, given what *S* knows at *w*. For *w* must be a possible situation, given what *S* knows.

I don't think Hossack's use of the notion of an epistemic alternative to demonstrate ignorance is entirely uncontroversial. If you think that knowledge is possible on the basis of (known) warrants which are inconclusive then you will not find the notion persuasive. Since then, there will always be a situation compatible with what *S* knows in terms of warrants but in which her knowledge claim is false. So the notion only seems to apply if you think that knowledge claims must be based on conclusive knowledge. But even then, it is not at all obvious that the relevant basing knowledge will always be available. We seem to be flirting with regress. It is likely that there will be cases where the knowledge that rules out an epistemic alternative is just the very knowledge in question. Consider perceptually based knowledge. If you are a naïve realist then you will take perception to give direct knowledge of the relevant state of affairs. There is no other knowledge one could appeal to to rule out epistemic alternatives. If not, you might think of perceptual knowledge as based on seemings which are compatible, given a multitude of scenarios involving deception, with the falsity of one's knowledge. But, if you thought this, you would be attracted to the notion of inconclusive epistemic warrants and, on that basis, reject the idea of epistemic

alternatives. So either the notion is inapplicable or potential cases can only be rejected trivially, that is, by appeal to the very knowledge in question. I think that it is plausible that knowledge acquired by stipulation is precisely of this sort. For it is brought into being by the act of stipulation and is not based on other pieces of knowledge. Moreover, I think my account of the mechanics of stipulation bears this out. Thus, I'm unconvinced by Hossack's invocation of epistemic alternatives. But I think it may be possible even to be concessive here and still find fault with the argument. Here's my concern.

Hossack concedes that:

S knows that in the language that S speaks 'Julius invented the zip' is true.

He explains:

S's stipulation has brought it about that S knows that S speaks some language or other in which the Julius sentence is true. It does not follow that the language *L* which S actually speaks is such that S knows that the Julius sentence is true in *L*. For in S's actual language 'Julius' denotes Julius and not Schmidt, so the Julius sentence 'Julius if anyone invented the zip' is true in S's actual language only if Julius invented the zip. But S has an epistemic alternative at which it was Schmidt not Julius who invented the zip, so S has an alternative at which the Julius sentence is false in S's actual language.

Hossack 2007: 154–155

Interestingly, Hossack uses the contingency of the Julius sentence to construct an epistemic alternative which undermines knowledge of it. The epistemic alternative, *w*, for S is one in which Schmidt, not Julius, invented the zip; that's what makes it relevantly alternative. So Schmidt isn't Julius and so, in the situation in which S makes her stipulation, Schmidt did not invent the zip. My thought is that *what S knows* is, indeed, incompatible with Schmidt being the inventor of the zip. So in *w*, a situation in which Schmidt invented the zip, S cannot have this knowledge. So *w* is not, *pace* Hossack, an epistemic alternative. Here's the argument spelled out:

(1) In the language that S speaks 'Julius invented the zip' is true.	Content of S's knowledge
(2) 'Julius invented the zip' is true in S's language iff the denotation of 'Julius' in S's language invented the zip.	Semantics of S's language & content of S's knowledge
(3) The denotation of 'Julius' in S's language invented the zip.	1 and 2, logic
(4) Schmidt invented the zip or did not invent the zip.	Logic
(5) If Schmidt invented the zip then Schmidt is the denotation of 'Julius'.	From 3
(6) If Schmidt is the denotation of 'Julius' then <i>w</i> is not an epistemic alternative.	Conceptual truth
(7) If Schmidt invented the zip then <i>w</i> is not an epistemic alternative.	From 5 and 6
(8) If Schmidt did not invent the zip then Schmidt is not the denotation of 'Julius'.	From 3

(9) If Schmidt did not invent the zip then w is incompatible with something S knows.	Since, in w , Schmidt is the denotation of 'Julius'; w is thus incompatible with 3, which follows by logic from knowledge that S has
(10) If w is incompatible with something S knows then w is not an epistemic alternative.	Conceptual truth
(11) If Schmidt did not invent the zip then w is not an epistemic alternative.	From 9 and 10
(12) w is not an epistemic alternative for S at w .	From 4, 7 and 11

Hossack is likely to complain that (2), being a disquotational fact, is only known to S in its wide scope version. But the reason he gives for this is that S has an epistemic alternative in which her stipulation brings it about that 'Julius' denotes Schmidt. But (2) doesn't *use* the name 'Julius' and the denotation of none of its terms is in doubt because of an epistemic alternative. The rest follows by means of logic from these two pieces of S's knowledge.

I anticipate another protest, perhaps more telling than this complaint, to be raised against the conditionals at (5) and (8), which have as antecedents 'Schmidt invented the zip' or its negation, respectively. If these are established by reasoning from the disjunct taken as a premise then we are reasoning from something which is neither something S knows, nor necessary, given what S knows. So the premise might be rejected as illegitimate in gauging whether a situation is an epistemic alternative. I think a good case can be made that the relevant conditionals are particular instances of generalized versions which hold by conceptual necessity. For instance, from 'The denotation of "Julius" in S's language invented the zip' it follows conceptually of any x that if x did not invent the zip then x is not the denotation in S's language of 'Julius'. So we just need to instantiate for x , and then use logic. The issue deserves more thought but will not be pursued any further; as I said earlier, though interesting, I don't think I need this argument.

4. Conclusions

1. The combination of entities in facts needs to accommodate the combination of entities in non-factual complexes or, equivalently, a whole variety of ways in which entities may occur in factual complexes – not simply as term or relating relation. The upshot is that there's no comprehending the ontology of facts.
2. The evidence in favour of construing knowledge as a relation to a fact is slight, even if the focus becomes knowledge-of rather than knowledge-that; since there seems to be little to distinguish knowledge from non-factive attitudes such as belief, except its factivity, and this doesn't justify the claim that they differ in logical form. So the relata of these relations might be the same, though they differ in their proper analysis, or in their semantics.
3. This conclusion about knowledge should be embraced, given the concerns about the ontology of facts raised in 1.

4. A *prioricity* is a property of pieces of knowledge-that; this, despite Hossack's urgings, is the issue the literature has traded in. So construed, there is good reason to distinguish it from necessity; and thus, little scope for explaining the latter modality in terms of knowledge. For, on the one hand, there appear to be *contents* that are necessary, though only knowable *a posteriori*; and on the other hand, our stipulative practice (or our ability to fix linguistic conventions) gives us epistemically unearned rights to assertions, whose content is not true of necessity.

Notes

- 1 For me, philosophical discussion with Keith is like a bracing swim in a choppy sea: you grapple or yield to the onrushing waves, but also feel the tug of deeper currents. In the main, I think I resist the currents, but am sure I've been enriched by the vigorous encounter with their force. My thanks to Keith for the swims, too few though they've been.
- 2 See (1903): 44–45.
- 3 See his *Philosophy of Logical Atomism*, lecture IV, reprinted as pp. 216–227 in his (1956) *Logic and Knowledge*.
- 4 I've argued (1995) that this is a difficult road to travel anyway on a Russellian construal of propositions because then, on the one hand, we want to say that the proposition is a unity because *love* appears as the relating relation; but, on the other hand, since *love* relates Desdemona to Cassio, it is hard to see how it can fail to be true.
- 5 Russell made the point in relation to facts which involve propositional attitude ascription as a result of a hard lesson he learned from Wittgenstein. Though the logic of propositional attitudes was central to his project, the same point applies to any case where we cannot treat the entities as a single complex entity, a fact, because the claim is false, and where there are two verbs, two relating relations. Thus, it applies to Hossack's account of negation and, I think, to other complexes – see below.
- 6 A new kind of combination – negation combination – indicated by boldface.
- 7 Equivalently it won't work for the fact, when it is a fact, of the negated conjunction.
- 8 We could make the same point as (ii) with negation but would need simply to compare facts. Suppose now $s \bullet (\neg, \textit{wisdom}, \textit{Donald})$ and $t \bullet (\neg, \textit{rarity}, \textit{beauty}, \textit{wisdom})$, then we have the same problem about distinguishing between the way *wisdom* occurs in these two facts, which now cannot be put down to its occupancy of the first position of the vector in *p* but not in *q*.
- 9 Or: S believes in God's goodness, though neither God nor her goodness exist.
- 10 Scare quotes because Hossack doesn't want to analyse knowledge.
- 11 Though I have queried the grounds for this.
- 12 Hossack rightly points out that 'Hesperus is Hesperus' is neither necessary nor *a priori*, since it entails that Hesperus exists. But he rightly points out we can retreat to the necessity and *a prioricity* of 'If Hesperus exists then Hesperus is Hesperus'. I'll stick with the inaccurate categorical identity statements; but everything I say could be recast to fit the conditional.
- 13 Compare. U asks, 'What is a vixen?', seeking an explanation of the meaning of 'vixen'. When told 'A vixen is a female fox', U learns a meaning and something *a priori*.

- 14 If U doesn't take the assertion to be true, then, arguably, she learns the meaning S attaches to 'Julius', but doesn't accept this meaning.
- 15 Another issue emerges about linguistic stipulation and/or convention. I think this is nicely brought out in Donnellan's and in Dummett's writings on this. Dummett (1973, appendix to ch. 5) downplays the phenomenon, without denying it, by showing how it is a product of a convention under which names are treated as always having wide scope relative to the modal operator. Donnellan (2012, ch. 6) asks how we are to know that the stipulation is to secure that the name is not being introduced as an abbreviation but as a rigid designator. If we do know that 'Julius' is a rigid designator or, equivalently, falls under a convention of being read with wide scope then we know *a priori* that Julius, if anyone, invented the zip; but also know (given the metaphysics of persons) that Julius might not have invented the zip. But, *if* we know that 'the inventor of the zip' is to be treated by convention as having wide scope (and why should we not choose to adopt this convention?), then similarly, we know *a priori* that the inventor of the zip, if anyone, invented the zip; but also know (given the metaphysics of persons) that the inventor of the zip might not have invented the zip. So, it seems to me, that focusing on the introduction of a name and on semantic descent is not relevant. The same points might be made in relation to conventions governing descriptions.
- 16 I think there are large questions here about whether the requisite notion is knowledge-of or knowledge-that, whether, that is we are thinking of *w* in terms of facts or of true contents.

References

- Donnellan, K. (2012). *Essays on Reference, Language and Mind*. J. Almog and P. Leonardi (eds). Oxford: Oxford University Press.
- Dummett, M. (1973). *Frege: Philosophy of Language*. London: Duckworth.
- Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Kitcher, P. (1980). 'Apriority and necessity'. *Australasian Journal of Philosophy* 58: 89–101.
- Russell, B. (1903). *Principles of Mathematics*. Cambridge: Cambridge University Press.
- Russell, B. (1956). *Logic and Knowledge*. R. C. Marsh (ed.). London: Unwin Hyman.
- Weiss, B. (1995). 'On the demise of Russell's multiple relations theory of judgement'. *Theoria: A Swedish Journal of Philosophy* 61(3): 261–282.

Necessity, Conditionals and Apriority

Keith Hossack

1. Inferribility

When we say ‘If A then C ’, what connection do we assert between the antecedent A and the consequent C ? Mill gives an inferribility account:

A hypothetical proposition is not . . . a mere aggregation of simple propositions. . . . Neither of these simple propositions may be true, and yet the truth of the hypothetical proposition may be indisputable. What is asserted is not the truth of either of the propositions, but the inferribility of the one from the other.

1967: 53

His view was endorsed by Ramsey:

In general we can say with Mill that ‘If p then q ’ means that q is inferrible from p , that is, of course, from p together with certain facts and laws not stated but in some way indicated by the context.

1929: 156

The Mill–Ramsey view is that a conditional ‘If A then C ’ is true if the antecedent A , possibly together with further relevant conditions S , (e.g. Ramsey’s ‘certain facts and laws’), entails the consequent C , i.e. if $A, S \vdash C$.

An inferribility account of the conditional can be premised on Hume’s reduction of causal necessity to regularity. According to Hume, what makes it true that ours is a world subject to laws of nature are the patterns of regularity in the world. Laws document regularities, and causal connection is a matter of inferribility from laws: A caused C if the occurrence of C can be inferred in the circumstances from the occurrence of A and the regular connection between A s and C s.

The Humean conception of causation gives rise to an inferribility account of the conditional. For example, scratching a match caused it to light. What is the causal explanation? According to Hempel (1965), if the laws and initial circumstances deductively entail the lighting of the match, then they causally explain it. Amongst the initial circumstances is the fact A that the match was scratched. Let S be the relevant

conditions, i.e. the laws L and the other initial circumstances. The causal explanation can be rewritten as $A, S \vdash C$. This says that C , the lighting of the match, is inferrible from the fact A that the match was scratched and the relevant conditions S . Thus if we know that relevant conditions S obtain, then if we know that $A, S \vdash C$, we can make a conditional prediction, that if the match is scratched, then it will light. Thus Hempel's account of explanation suggests that the conditional 'If A then C ' is true if there obtain relevant conditions S such that $A, S \vdash C$.

This approach can be naturally extended to handle a 'counterfactual' conditional, 'If it were that A , then it would be that C '. The counterfactual can be known to be true even when A is false. Consider a match that was never actually scratched. Suppose it had been scratched – would it have lit? Given what we know about matches, we suppose that scratching would indeed have caused it to light. On the inferribility conception of causation, we must postulate an entailment $A, S \vdash C$ between the scratching A , the lighting C , and the (now hypothetical) relevant circumstances S . This entailment makes 'If A then C ' true even in the counterfactual case.

Goodman examined whether the entailment idea can be spelled out in sufficient detail to give an analysis of counterfactuals. He suggested that a counterfactual 'If A then C ' is true if and only if C is inferrible either from A alone, or from A together with other 'relevant conditions' S . This is just what we are led to, given the Humean conception of causal explanation. Goodman noted three necessary conditions if S is to be 'relevant' in the sense required (1979: 10–13).

1. A and S must be consistent.
2. If S are relevant conditions and $A, S \vdash C$, there must not exist relevant S such that $A, S \vdash \neg C$.
3. The relevant circumstances S must be 'cotenable' with the antecedent A : by this Goodman meant that it must not be the case that $A \Box \rightarrow \neg S$.

The reason Goodman needed his third condition of cotenability was to block certain troublesome counterexamples. He writes:

The trouble is caused by including in our S a statement that would not have been true if A were. Accordingly we must exclude such statements from the set of relevant conditions; S , in addition to satisfying the other requirements already laid down, must be not merely compatible with A but 'jointly tenable' or *cotenable* with A . A is cotenable with $S \dots$ if it is not the case that S would not be true if A were.

1979: 8

Goodman's initial optimism about the possibility of analysing counterfactuals ended with his discovery of the cotenability condition, since the definition of cotenability itself requires the counterfactual. He writes:

Thus we find ourselves in an infinite regressus or a circle; for cotenability is defined in terms of counterfactuals, yet the meaning of counterfactuals is defined in terms of cotenability. In other words to establish any counterfactual, it seems that we first

have to determine the truth of another. If so, we can never explain a counterfactual except in terms of others, so that the problem of counterfactuals must remain unsolved.

1979: 16–17

Because Goodman was unable to define cotenability, and hence unable to define ‘relevant circumstances’, the truth-conditions he gave are incomplete and do not help to clarify the logic of conditionals. In this respect his account is inferior to the accounts of Stalnaker and Lewis, who propose truth-conditions that turn on the idea of the far-fetchedness of a possible world. According to Stalnaker (1968), $A \Box \rightarrow C$ is true if and only if C is true at the least far-fetched world at which A is true. According to Lewis (1973a), $A \Box \rightarrow C$ is true if and only if some world at which A and C are both true is less far-fetched than any world at which A is true but C is false.

Stalnaker and Lewis are right that far-fetchedness is key to the logic of counterfactuals. But we can incorporate far-fetchedness in Goodman’s account of the conditional, without bringing in possible worlds, by defining his ‘relevant’ as follows:

Definition. A proposition S is said to be *relevant to A* if S and A are consistent, and every proposition consistent with S is less far-fetched than any proposition inconsistent with S .

Given this definition of ‘relevant’, Theorem 1 of the Appendix proves that if S is relevant to A then:

- (i) A and S are consistent;
- (ii) if $A, S \vdash C$, there is no relevant S such that $A, S \vdash C$;
- (iii) A and S are cotenable.

Thus this definition of relevance meets the three conditions that Goodman identified.

2. Far-fetchedness

Our proposed completion of Goodman’s truth-conditions defines ‘relevance’ in terms of the relative far-fetchedness of propositions. But what is the ‘far-fetchedness’ of a proposition? Lewis sought to analyse it in terms of the comparative overall similarity of possible worlds: P is less far-fetched than Q if P is true at some world more similar overall to the actual world than any world at which Q is true (1979: 465). According to Lewis, the fruitfulness of this reduction speaks in favour of his hypothesis of a ‘plurality of worlds’, since the worlds must really exist if their comparative similarity is to be an objective matter.

It may be doubted that this reduction is entirely successful, for the appeal to ‘similarity’ requires the conditional to elucidate it. For what are the respects of similarity that matter in judgements of overall similarity of worlds? Responding to an objection by Kit Fine (1975: 452), Lewis says we must use our knowledge of conditionals to identify the relevant respects of similarity:

We must use what we know about the truth and falsity of counterfactuals to see if we can find some sort of similarity relation ... to yield the proper truth conditions. ... we must use what we know about counterfactuals to find out about the appropriate similarity relation—not the other way round.

1979: 467

As Lewis continues his discussion, it becomes clear that he intends to rely on ‘what we know about counterfactuals’ to settle the ‘appropriate’ respects of similarity. So it seems that his reductive analysis of far-fetchedness presupposes the counterfactual – we could not understand what he means by ‘similarity’ unless we already understood the counterfactual.

However, a reductive analysis of far-fetchedness is not required. We can make intuitive judgements directly about how ‘far-fetched’ a proposition is, without needing a detour through possible worlds. For example, intuitively it is far-fetched that a donkey should talk, further-fetched that it should talk philosophy, and further-fetched still that it should talk philosophy with a unicorn. We can express far-fetchedness in terms of the conditional, as follows: P is less far-fetched than Q if, even were it the case that P , still it would not be the case that Q . For example, even if a donkey talked, still it would not talk philosophy: and even if a donkey talked philosophy, still it would not talk philosophy with a unicorn.

We need not rely only on intuitive judgements, for far-fetchedness can be defined explicitly in terms of the counterfactual, as follows:

Definition P is said to be *less far-fetched* than Q , written ‘ $P < Q$ ’ if, were exactly one of P and Q the case, then P would be the case and Q would not be the case.

In symbols: $P < Q =_{\text{Df}} \neg(P \leftrightarrow Q) \Box \rightarrow (P \wedge \neg Q)$ ¹

We were hoping to complete Goodman’s truth-conditions for the counterfactual by defining ‘relevant’ in terms of far-fetchedness. But now we are defining ‘far-fetchedness’ in terms of the counterfactual. Are we going round in circles here? We would be, if with Goodman we were trying to give a non-circular analysis of the counterfactual conditional. But all we are now inquiring about is truth-conditions. It is customary to give the truth-conditions of logical constants by reuse of the constants themselves in the metalanguage. The correctness of such truth-conditions is not impugned by their failure to give a non-circular definition of the logical constant concerned – no such definition is needed, or indeed possible. So it would not be an objection to Goodman’s truth-conditions, if they required the counterfactual in the metalanguage. On the contrary, that would fit well with the view that ‘if’ is an indefinable logical primitive, as are ‘and’, ‘not’ and ‘all’.

3. Relative correctness of the truth-conditions of Goodman and of Lewis

We have defined ‘relevant’ in such a way as to connect Goodman’s truth-conditions with the truth-conditions of Stalnaker and Lewis. To study the connection further, we need the following chain of definitions. If Γ is any class of propositions, then:

Definition Γ is said to be *complete* if given any proposition A , $\Gamma \vdash A$ or $\Gamma \vdash \neg A$.

Definition Γ is said to be *consistent* if given any proposition A , not both $\Gamma \vdash A$ and $\Gamma \vdash \neg A$.

Definition The proposition that every proposition in Γ is true is called the *infinite conjunction* of Γ , written ' $\cap\Gamma$ '.

Definition The proposition that some proposition in Γ is true is called the *infinite disjunction* of Γ , written ' $\cup\Gamma$ '.

Definition Γ is said to be an *encyclopaedia* if Γ is complete and consistent.

Definition A proposition $\cap\Gamma$ is called a *history* if Γ is an encyclopaedia.

The truth-conditions of Stalnaker and Lewis turn on the concept of truth at a possible world. For each world w , we can imagine an encyclopaedic class of propositions that tells the whole truth about how things would be if w were actual. The class can be visualized as the limit of a process of adding more and more consistent detail to some initial partial description of the world. It will be seen that for each possible world, there is exactly one encyclopaedia. There is at least one, because for every world w and proposition A , either A or $\neg A$ is true at w : so for every proposition A , either A or $\neg A$ is in the encyclopaedia. And there is at most one, for if Γ and Γ' are different then for some proposition A , Γ contains A but Γ' contains $\neg A$, so A and $\neg A$ are both true at w , which is absurd (Plantinga 1974: 46).

Thus for every world w there exists an encyclopaedia Γ . Whether the converse is true depends on how abundant the possible worlds are. On Stalnaker's semantics worlds are not very abundant, since on his view we need only enough worlds to deliver a theory of the content of thought. But Lewis thinks, surely rightly, that some possibilities are too complex for human thought to distinguish, so for him worlds cut much finer than thoughts. For Lewis, therefore, worlds must exist in formidable abundance. His 'Principle of Plenitude' (1986: 86–92) says there are as many worlds as there are ways of populating a spacetime with instantiations of perfectly natural properties. Since it is none too clear which properties are 'perfectly natural', or what exactly counts as a 'spacetime', this leaves the exact abundance of worlds, and hence also the exact logic of counterfactuals, somewhat indeterminate. But we can remove the indeterminacy by postulating, following Plantinga, that *every* complete consistent class of propositions is the encyclopaedia of some possible world.

Plantinga's Postulate For every complete consistent class Γ of propositions there exists a possible world w such that Γ is the encyclopaedia of w .

1974: 46

Theorem 2 of the Appendix proves that, given Plantinga's Postulate, Goodman's truth-conditions for the counterfactual are equivalent to Lewis's. So Goodman's truth-conditions are correct relative to Lewis's, provided possible worlds are precisely as abundant as Plantinga's Postulate says they are.

5. Modal holism

Leibniz's theory of 'possible universes' (1991: §53) is an extremely helpful logical tool. Almost all contemporary philosophers freely use the theory in the investigation of arguments of any intricacy, yet very few believe that there really are such entities as possible worlds. As Rosen (1990) observes, most treat possible worlds as a useful fiction. Of course a fiction doesn't have to be true to be good, and what is so very good about this particular fiction is this: if we reason about possibilities as if the fiction were true, our modal reasoning will be accurate. Even if there is no fact of the matter about whether possible worlds really exist, there certainly is a fact of the matter about whether our modal reasoning is correct. So by mining what is implicit in the very idea of possible worlds, we should be able to discover important principles of the logic of possibility.

When we are considering whether a proposition is actually true, holism is required: no proposition faces the tribunal of reason on its own, but only as part of some bigger picture that is to be assessed as a whole, all things considered. The fundamental insight implicit in the idea of possible worlds is that this holism applies also when we are considering whether a proposition is possibly true: that too has to be considered as part of a bigger picture. The idea of possible worlds enforces this holism by saying that a proposition is possible only if there is a whole world at which the proposition is true. According to Adams, we thus arrive at 'the intuitively very plausible thesis that possibility is holistic rather than atomistic, in the sense that what is possible is possible only as part of a possible completely determinate world' (1974: 225). This 'intuition' can be given formal expression as follows:

The Holism Axiom If a proposition A is possible, there exists a complete class of propositions the conjunction of which is possible and which contains A .

According to Stalnaker, a thesis of this sort 'assumes nothing about propositions except what is required to capture the principle intuition of the theory: that something is possible only as part of a possible completely determinate situation' (1984: 53, 55).

The holism axiom is implicit in the theory of possible worlds. For suppose the proposition A is possible. Then according to the theory of possible worlds, there is a possible world w such that A is true at w . Let Γ be the class of all propositions true at w . Then $A \in \Gamma$. Also, for every proposition B , either B or $\neg B$ is in Γ , since either B or $\neg B$ is true at w . So Γ is complete. Since every proposition in Γ is true at w , the infinite conjunction $\bigcap \Gamma$ is true at w , and hence possible. If the holism axiom is true, its implicit presence in the theory of possible worlds explains why use of the theory is reliable in modal reasoning, even if the worlds themselves are fictions. I therefore conclude that the holism axiom is true: it is used in the proof of Theorem 3 below.

6. Absolute correctness of the truth-conditions of Goodman

Plantinga's Postulate sets up a one-one correspondence between possible worlds and their encyclopaedias, and hence between worlds and their histories. The proof of Theorem 2 in the Appendix exploits this correspondence. However, given the existence

of the one-one correspondence, we don't need both worlds and histories for the semantics. Lewis chose to do without histories and to rely solely on his ontology of worlds as parallel universes. An alternative is to use the holism axiom to do without worlds, and then use histories to prove outright that Goodman's truth-conditions are correct, without mentioning possible worlds at all.

We can do this by using in the metalanguage the known principles of the logic of counterfactuals. This is the same technique that is standardly used to prove the correctness of truth-conditions for other logical constants. Given a proposed semantics, we prove its correctness by using in the metatheory the known axioms and rules of inference governing the logical constant in question. Similarly, for the logical constant 'if', Goodman's truth-conditions can be proved correct by means of the following eight intuitively evident axioms of the logic of counterfactuals.

1. If A is consistent, then $A \Box \rightarrow C \leftrightarrow (A \wedge C) < (A \wedge \neg C)$
(If A is consistent, then $A \Box \rightarrow C$ is true if and only if $A \wedge C$ is less far-fetched than $A \wedge \neg C$.)
2. If A is inconsistent, then $A \Box \rightarrow C$ is true.
3. If A is consistent and B is inconsistent, then A is less far-fetched than B .
4. *Transitivity* $A < B \wedge B < C \rightarrow A < C$
(If A is less far-fetched than B , and B is less far-fetched than C , then A is less far-fetched than C .)
5. *Asymmetry* $A < B \rightarrow \neg(B < A)$
(If A is less far-fetched than B , then B is not less far-fetched than A .)
6. If $\cup \Gamma < A$, then there exists some $g \in \Gamma$ such that $g < A$.
(If it is less far-fetched that some proposition in Γ is true than that A is true, then some proposition in Γ is less far-fetched than A .)

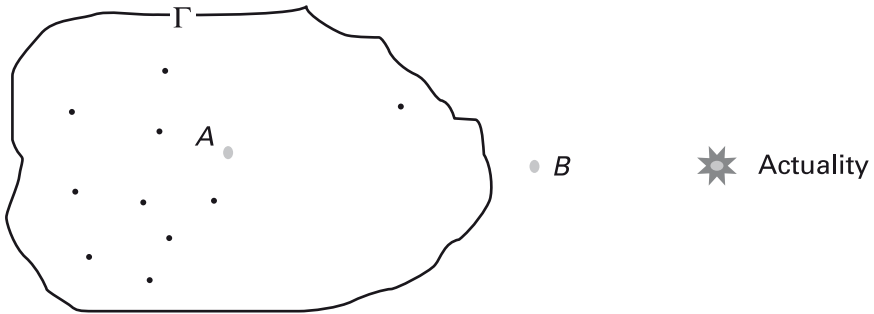


Figure 7.1 Comparative far-fetchedness of class and element.

7. $B < \cup \Gamma$ if and only if $g \in \Gamma \rightarrow B < g$.
(The following are equivalent: (i) it is less far-fetched that B than that some proposition in Γ is true; (ii) B is less far-fetched than every proposition in Γ .)
8. *Interchange of logical equivalents* If A is logically equivalent to A , and B is logically equivalent to B , then A is less far-fetched than B if and only if A is less far-fetched than B .

Theorem 3 of the Appendix proves that these axioms entail that a counterfactual conditional $A \Box \rightarrow C$ is true if and only if it satisfies Goodman's truth-conditions. Therefore, Goodman has indeed given us exactly the right truth-conditions for counterfactuals.

7. Defining necessity in terms of the counterfactual

According to Leibniz, what distinguishes the human power of reason from the intellectual capacities of simpler creatures is our knowledge of necessary truths:

§29. But the knowledge of necessary and eternal truths is what distinguishes us from mere animals and provides us with *reason* and the sciences, elevating us to a knowledge of ourselves and of God.

1991: 20

Leibniz equated the necessary truths with the *a priori* truths of reasoning:

§33 There are two kinds of truth, those *of reasoning* and those *of fact*. Truths of reasoning are necessary and their opposite is impossible, while those of fact are contingent and their opposite is possible. When a truth is necessary one can find its reason through analysis, resolving it into ever more simple ideas and truths until one reaches primitives.

1991: 21

Goodman's truth-conditions make the counterfactual an intermediary between necessity and the *a priori*, for necessity can be defined in terms of the counterfactual, and the counterfactual can be defined in terms of the *a priori*. This allows a proof that Leibniz's modal rationalism is correct: the necessary is indeed the *a priori*.

The first step in this proof is to define necessity in terms of the counterfactual, thus reversing the usual process of defining the counterfactual in terms of possible worlds. A counterfactual definition was given by Hume, who says that a proposition is necessary if the contrary 'implies a contradiction' (1975: §21). The same definition is given by Lewis (1973b: 426), who symbolizes it thus: $\Box A =_{\text{df}} \neg A \Box \rightarrow \perp$.

Hume's definition. A proposition C is called *necessary* if, were $\neg C$ the case, a contradiction would be the case.

Mill defines the necessary as that which would be true, whatever else were the case:

If there be any meaning which confessedly belongs to the term necessity, it is *unconditionality*. That which is necessary, that which *must* be, means that which will be, whatever supposition we may make in regard to all other things.

1967: 222

Mill's definition. A proposition C is called *necessary* if, whatever else were the case, C would be the case.

A particularly simple definition is given by Williamson:

Simple definition. A proposition C is called *necessary* if, were $\neg C$ the case, C would be the case.

2007: 155–161

That the three definitions are equivalent can be checked as follows.

Mill entails Hume. Assume that C is necessary on Mill's definition. Then whatever other proposition A may be, if it were that A it would be that C . In particular, if it were that $\neg C$ it would be that C , and hence $C \wedge \neg C$ would be true. Thus if it were that $\neg C$ a contradiction would be true. So C is necessary on Hume's definition.

Hume entails simple. Assume that C is necessary on Hume's definition. So if it were that $\neg C$ then a contradiction would be true. But if a contradiction were true, everything would be true, so in particular C would be true. Therefore if it were that $\neg C$, it would be that C . So C is necessary on the simple definition.

Simple entails Mill. Assume that C is necessary on the simple definition. So if $\neg C$ were the case, C would be the case also. So if $\neg C$ were the case, then the contradiction $C \wedge \neg C$ would be the case. But it is impossible that a contradiction should be the case, so it is impossible that $\neg C$. So for any other proposition A , if it were that A it would not be that $\neg C$, so if it were that A it would be that C . So C is necessary on Mill's definition.

The agreement of these three definitions confirms that necessity can indeed be defined in terms of the counterfactual.

8. Defining the *a priori*

How should we define '*a priori*'? According to Leibniz, a proposition is *a priori* when it has a proof, i.e. if 'one can find its reason through analysis, resolving it into ever more simple ideas and truths until one reaches primitives' (1991: 21). Frege says the *a priori* truths are those that have a deductive proof from primitive truths (1978: §3). The Leibniz–Frege definition is that a proposition is *a priori* if it is primitively evident or a logical consequence of primitively evident truths.

The thesis that necessity is apriority allows us to reduce the modal to the epistemic. The reduction would fail if the definition of '*a priori*' concealed an unreduced modal notion, but a definition in the spirit of Leibniz and Frege does not do so. The primitively evident is that which is disposed to be known immediately by the faculty of reason: this definition is epistemic and causal, but not modal. Logical consequence is also an epistemic notion, for C is a logical consequence of a class of premisses Σ if C is inferrible

from Σ , i.e. if there is a proof of C from Σ . A proof is defined as a chain of propositions each of which is either primitively evident, or one of Σ , or which stands to preceding propositions in a relation that is evidently truth-preserving. I will write ' $\Sigma \vdash C$ ' to mean there is a proof of C from Σ . If C can be proved without the need for any premisses I write ' $\vdash C$ ', and define the *a priori* accordingly:

Definition. A proposition C is said to be *a priori* if $\vdash C$.

This definition does not imply that every *a priori* proposition is humanly knowable. A proof is a chain of propositions, which may be infinite, if for example the proof uses the ω -rule or the rule of Complete Induction (Hossack 2020: 31–32). Even if the chain of propositions is finite it may be too long to be humanly taken in, so the chain as a whole may not be surveyable. Moreover, a truth-preserving relation can obtain between links in a chain whether or not we can know the relation obtains. Finally, a proposition may be *a priori* but not knowable, if no one will ever entertain the proposition.

9. The truth-conditions of Goodman entail modal rationalism

On these definitions, the truth-conditions of Goodman entail that necessity and apriority coincide. The proof of this is as follows. According to Goodman's truth-conditions, a counterfactual conditional $A \Box \rightarrow C$ is true if and only if either (i) $A \vdash C$, or (ii) there exist 'relevant conditions' S such that $A, S \vdash C$. So suppose C is *a priori*. Then $\vdash C$, so $\neg C \vdash C$, so $\neg C \Box \rightarrow C$, by Goodman's truth-condition (i), so C is necessary, by the simple definition of necessity. For the converse, suppose C is necessary. Then $\neg C \Box \rightarrow C$ by the simple definition. So by Goodman's truth-conditions, either (i) $\neg C \vdash C$ or (ii) there is a relevant S such that $\neg C, S \vdash C$. But the latter is impossible. For if $\neg C, S \vdash C$, then $\neg C \vdash \neg S$, so $\neg C \Box \rightarrow \neg S$ by truth-condition (i), so S is not cotenable with $\neg C$, so S is not relevant to $\neg C$, by Theorem 1 of the Appendix. So truth-condition (ii) is not met. Therefore truth-condition (i) is met, so $\neg C \vdash C$, so $\vdash C$, i.e. C is *a priori*.

Thus if Goodman's truth conditions are correct, then the necessary is the *a priori*. But Theorem 3 of the Appendix proves that Goodman's truth conditions are correct. Therefore, the necessary is indeed the *a priori*.

Appendix – Proofs of theorems

Lemma 1. $S \vdash T$ if and only if every proposition inconsistent with T is inconsistent with S .

Proof. Assume that every proposition inconsistent with T is inconsistent with S . Then since $\neg T$ is inconsistent with T , it is inconsistent with S , so $S, \neg T \vdash \perp$, so $S \vdash T$. Conversely, assume $S \vdash T$. If P is inconsistent with T , then $P, T \vdash \perp$, so $P, S \vdash \perp$, so P is inconsistent with S .

Lemma 2. If S and T are both possible, then if S is less far-fetched than T , then T is not less far-fetched than S .

Proof. Suppose that S is less far-fetched than T , and that T is less far-fetched than S . Then because S is less far-fetched than T , if either S or T were the case, S would be the case but T would not be the case. But because T is less far-fetched than S , if either T or S were the case, T would be the case but S would not be the case. So if either S or T were the case, a contradiction would be true, so neither S nor T is possible, contrary to our assumption.

Lemma 3. If S and T are circumstances relevant to A , then $S \vdash T$ or $T \vdash S$.

Proof. If $S \vdash T$ there is nothing to prove. Otherwise by Lemma 1, some proposition P is inconsistent with T but consistent with S . Let Q be any proposition, and suppose Q is inconsistent with S . Assume for *reductio* that Q is consistent with T . Then Q and T are consistent, but P and T are inconsistent; so by the relevance of T , Q is less far-fetched than P . But P and S are consistent, and Q and S inconsistent; so by the relevance of S , P is less far-fetched than Q , contradicting Lemma 2. So Q is inconsistent with T , so $T \vdash S$, by Lemma 1.

Theorem 1. If S is relevant to A then (i) A and S are consistent; (ii) if $A, S \vdash C$, there is no relevant S such that $A, S \vdash C$; (iii) A and S are cotenable.

Proof. (i) is immediate from our definition of 'relevant'. For (ii), assume S is relevant to A and that $A, S \vdash C$. Assume for *reductio* that there exists S relevant to A such that $A, S \vdash \neg C$. By Lemma 3, either $S \vdash S$ or $S \vdash S$. If $S \vdash S$, then since $A, S \vdash \neg C$, it follows that $A, S \vdash \neg C$; but $A, S \vdash C$, so A and S are inconsistent, contrary to the relevance of S . So $S \vdash S$. Since $A, S \vdash C$, it follows that $A, S \vdash C$. But $A, S \vdash \neg C$, so A and S are inconsistent, contrary to the relevance of S . Contradiction. So there is no relevant S such that $A, S \vdash C$.

For (iii), assume S is relevant to A . Assume for *reductio* that A and S are not cotenable, so $A \Box \rightarrow \neg S$. Since A and S are consistent by the relevance of S , we may rule out the possibility that $A \vdash \neg S$. Hence there exists some T relevant to A such $A, T \vdash \neg S$. Since S and T are both relevant, $S \vdash T$ or $T \vdash S$, by Lemma 3. T is relevant to A , and $A, T \vdash \neg S$, so we may rule out the case $T \vdash S$, since then $A, T \vdash S$, contrary to the consistency of A and T . So $S \vdash T$, so $A, S \vdash T$. But $A, T \vdash \neg S$, so $A, S \vdash \neg T$. Therefore A and S are inconsistent, so S is not relevant to A . Contradiction. So A and S are cotenable.

Lemma 4. $\Diamond A \leftrightarrow A < \perp$

(A proposition is possible if and only if it is less far-fetched than a contradiction.)

Proof. Assume $\Diamond A$. Suppose $A \prec \perp$. Then A is inconsistent by [3]. (Numbers in square brackets here refer to the axioms listed in Section 6.) So $A \Box \rightarrow \perp$ by [2]. So $\neg \Diamond A$ by Hume's definition. Contradiction. For the converse, assume $A < \perp$. Suppose A is not possible. Then by Hume, $A \Box \rightarrow \perp$. So $A \wedge \perp < A \wedge \neg \perp$ by [1], so $\perp < A$ by [8], contradicting [5].

Lemma 5. Every consistent class of propositions has a complete consistent extension.

Proof. Let Γ be a consistent class of propositions. Then $\cap\Gamma$ is consistent. So $\cap\Gamma$ is less far-fetched than $\perp$, by [3]. So $\cap\Gamma$ is possible by Lemma 4. So by the Holism Axiom of section 5 there is a complete class of propositions Δ such that $\cap\Gamma \in \Delta$ and $\cap\Delta$ is possible. Since $\cap\Delta$ is possible, $\cap\Delta < \perp$ by Lemma 4, so $\cap\Delta$ is consistent, so Δ is consistent. If $p \in \Gamma$ then $\cap\Gamma \vdash p$, so $\Delta \vdash p$, so $p \in \Delta$, so $\Gamma \subseteq \Delta$.

Theorem 2. Plantinga's Postulate entails that a counterfactual is true according to Lewis if and only if it is true according to Goodman.

Proof. Suppose the counterfactual $A \Box \rightarrow C$ is true according to Lewis. Then either there are no A -worlds, or some $(A \wedge C)$ -world is less far-fetched than any $(A \wedge \neg C)$ -world. If there are no A -worlds then, by Plantinga's Postulate, no complete consistent class of sentences contains A . So by Lemma 5, $\{A\}$ is inconsistent, so $\vdash \neg A$, so $A \vdash C$, so $A \Box \rightarrow C$ is true according to Goodman. But if there are A -worlds, let w be an $(A \wedge C)$ -world that is less far-fetched than every $(A \wedge \neg C)$ -world. Let Σ be the class of all worlds no further-fetched than w , and let X be the class of encyclopaedias of the worlds in Σ . Let $S = \bigcup \cap X$ be the infinite disjunction of the infinite conjunctions of X : then S is the infinite disjunction of all the histories of all the worlds no more far-fetched than w . Since $A \wedge \neg C$ is true only at worlds more far-fetched than w , $A \vdash C$ is true at every world in Σ . So for every Γ in X , $A \vdash C \in \Gamma$. So $S \vdash A \vdash C$. So $S, A \vdash C$. We must prove that S is relevant to A . Since S and $A \wedge C$ are both true at w , S and A are consistent. Moreover, let P be any proposition consistent with S , and let Q be any proposition inconsistent with S . Then since $S = \bigcup \cap X$, P can be consistent with S only if for some $\Gamma \in X$, P is consistent with Γ , so $\Gamma \not\vdash \neg P$, so since Γ is complete, $\Gamma \vdash P$. So by Plantinga's Postulate, there is a world w' whose encyclopaedia is Γ . So P and S are both true at w' ; similarly, since Q and S are inconsistent, S is false at every world w' at which Q is true. Since worlds at which S is true are less far-fetched than worlds at which S is false, it follows that P is true at a world w' that is less far-fetched than any world w' at which Q is true, i.e. P is less far-fetched than Q . So S is relevant to A . So $A \Box \rightarrow C$ is true according to Goodman.

For the converse, suppose $A \Box \rightarrow C$ is true according to Goodman. Then either $A \vdash C$, or there is a relevant S such that $A, S \vdash C$. Suppose $A \vdash C$. Then any consistent complete extension of $\{A\}$ contains C , so any history that entails A entails C , so either there are no A -histories, and hence by Plantinga's Postulate there are no A -worlds; or trivially some $(A \wedge C)$ -world is less far-fetched than any $(A \wedge \neg C)$ -world, since there are none of the latter; so $A \Box \rightarrow C$ is true according to Lewis. But if $A \not\vdash C$, there is a relevant S such that $A, S \vdash C$. Then since A and S are consistent, the proposition $A \wedge C$ is consistent with S , but the proposition $A \wedge \neg C$ is inconsistent with S . So by the relevance of S , the proposition $A \wedge C$ is less far-fetched than the proposition $A \wedge \neg C$; so according to Lewis, some $(A \wedge C)$ -world is less far-fetched than any $(A \wedge \neg C)$ -world, i.e., $A \Box \rightarrow C$ is true according to Lewis. So if $A \Box \rightarrow C$ is true according to Goodman, it is true according to Lewis.

Theorem 3. A counterfactual conditional $A \Box \rightarrow C$ is true if and only if it satisfies Goodman's truth-conditions.

Proof. Suppose $A \Box \rightarrow C$ is true. We must prove that either $A \vdash C$, or that there exists a relevant S such that $A, S \vdash C$. If $\{A\}$ is inconsistent or if $\{A, \neg C\}$ is inconsistent, then $A \vdash C$. Otherwise $\{A, \neg C\}$ is consistent, so $A \wedge \neg C$ is consistent. Since $A \Box \rightarrow C$ is true, $(A \wedge C) < (A \wedge \neg C)$ by [1]. (Numbers in square brackets again refer to the axioms listed in Section 6.) So $A \wedge C$ cannot be inconsistent, by [3]. Let Γ be the class of all histories that entail $A \wedge C$, and let Δ be the class of all histories that entail $A \wedge \neg C$. Then since $\{A, C\}$ and $\{A, \neg C\}$ are consistent, each has a consistent complete extension, by Lemma 5, so neither Γ nor Δ is empty. $\cup \Gamma$ entails that some history $h \in \Gamma$ is true, and every history in Γ entails $A \wedge C$, so $\cup \Gamma \vdash A \wedge C$. Since $A \wedge C \vdash \cup \Gamma$, therefore $\cup \Gamma$ is logically equivalent to $A \wedge C$. Similarly, $\cup \Delta$ is logically equivalent to $A \wedge \neg C$. So since $(A \wedge C) < (A \wedge \neg C)$, it follows by interchange of logical equivalents [8] that $\cup \Gamma < \cup \Delta$. So by [6], there is some $g \in \Gamma$ such that $g < \cup \Delta$. So by [7], for every $d \in \Delta$, $g < d$. Let Σ be the class whose elements are histories h such that $h \leq g$. Suppose Σ and Δ have a common member h . Since $h \in \Sigma$, $h \leq g$. Since $h \in \Delta$, $g < h$, contradicting [5]. So Δ and Σ have no common member, so every history in Δ is inconsistent with every history in Σ , so $\cup \Delta$ and $\cup \Sigma$ are inconsistent. Let S be the disjunction $\cup \Sigma$ of all the histories in Σ . We prove that S is a relevant proposition such that $A, S \vdash C$. Since S and $A \wedge \neg C$ are inconsistent, $A, S \vdash C$. It remains to prove that S is relevant to A . Now g is consistent, so since $g \vdash S$ and $g \vdash A \wedge C$, it follows that S and A are consistent. Moreover, let P be any proposition consistent with S , and let Q be any proposition inconsistent with S . Since P and S are consistent, P is consistent with some history h in Σ , so h is a history in which P and S are both true; similarly, since Q and S are inconsistent, S is false in every history in which Q is true. Let Π be the class of all histories according to which P is true, and let Θ be the class of all histories according to which Q is true. If Θ is empty then Q is inconsistent, so since P is consistent, $P < Q$ by [3]. Otherwise, Q is consistent. P and $\cup \Pi$ are logically equivalent, and Q and $\cup \Theta$ are logically equivalent, so to prove that $P < Q$, it suffices to prove that $\cup \Pi < \cup \Theta$. Choose $p \in \Pi$. Then p is a history and hence complete, so either $p \vdash S$ or $p \vdash \neg S$. But since P and S are consistent, $p \not\vdash \neg S$, so $p \vdash S$, so $p \in \Sigma$, so $p \leq g$. So for every $p \in \Pi$, $p \leq g$, so $\cup \Pi \leq g$ by [7], so $P \leq g$. Let $q \in \Theta$. Since $Q = \cup \Theta$ is inconsistent with S , $q \vdash \neg S$, so $q \not\in \Sigma$. So for every $q \in \Theta$, $g < q$, so $g < \cup \Theta$ by [7]. So $g < Q$. Since $P \leq g$ and $g < Q$, $P < Q$, by [4]. So S is relevant. So $A \Box \rightarrow C$ satisfies Goodman's truth-conditions

For the converse, suppose that $A \Box \rightarrow C$ satisfies Goodman's truth-conditions. Then if $A \vdash C$ then if A is inconsistent then $A \Box \rightarrow C$ is true, by [2]. Otherwise $\{A, C\}$ is consistent, so since $A \vdash C$, $\{A, \neg C\}$ is inconsistent, so $(A \wedge C) < (A \wedge \neg C)$ by [3], so $\neg [(A \wedge C) \leftrightarrow (A \wedge \neg C)] \Box \rightarrow [(A \wedge C) \wedge \neg(A \wedge \neg C)]$ by the definition of far-fetchedness. But $\neg [(A \wedge C) \leftrightarrow (A \wedge \neg C)]$ is logically equivalent to A , and $[(A \wedge C) \wedge \neg(A \wedge \neg C)]$ is logically equivalent to C , so in consequence of [7], $A \Box \rightarrow C$ is true. But if $A \not\vdash C$, then there is some relevant proposition S such that $A, S \vdash C$. Then $A \wedge C$ is consistent with S , but $A \wedge \neg C$ is inconsistent with S . So by the relevance of S , $(A \wedge C) < (A \wedge \neg C)$, so $A \Box \rightarrow C$ is true by [1]. So if $A \Box \rightarrow C$ satisfies Goodman's truth-conditions, then $A \Box \rightarrow C$ is true.

Note

- 1 *Endnote added in proof stage.* This definition is in error, as Dorothy Edgington points out in her 'Commentary' paper in this volume. To correct the definition, the following should be substituted:

Definition P is said to be less far-fetched than Q , written ' $P < Q$ ', if P is consistent and if P or Q were the case, then Q would not be the case.

(In symbols: $P < Q =_{\text{Df}} P$ is consistent and $(P \vee Q) \Box \rightarrow \neg Q$.)

Consequently the second and third sentences of the last paragraph of the proof of **Theorem 3** should be replaced with the following: 'Otherwise $\{A, C\}$ is consistent, so since $A \vdash C$, $\{A, \neg C\}$ is inconsistent, so $(A \wedge C) < (A \wedge \neg C)$ by (3), so by (1), $A \Box \rightarrow C$ is true.'

References

- Adams, R. M. (1974). 'Theories of actuality', *Noûs* 8: 211–231.
- Fine, K. (1975). 'Review of Lewis (1973a)', *Mind* 84: 451–458.
- Frege, G. (1978). *The Foundations of Arithmetic*, trans. J. L. Austin, Evanston, IL: Northwestern University Press.
- Goodman, N. (1979). 'The problem of counterfactual conditionals', in his *Fact, Fiction and Forecast*. Cambridge MA: Harvard University Press, pp. 3–27.
- Hempel, C. (1965). 'Aspects of scientific explanation', in his *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number*. London: Bloomsbury.
- Hume, D. (1965). *A Treatise of Human Nature*, ed. L. Selby-Bigge. Oxford: Oxford University Press.
- Hume, D. (1975). 'An enquiry concerning human understanding'. In L. Selby-Bigge (ed.), *Hume's Enquiries*. Oxford: Oxford University Press.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
- Leibniz, G. (1991). *Monadology*, translated in N. Rescher, G. W. *Leibniz's Monadology*. London: Routledge, pp. 17–29.
- Lewis, D. (1973a). *Counterfactuals*. Oxford: Basil Blackwell.
- Lewis, D. (1973b). 'Counterfactuals and comparative possibility', *Journal of Philosophical Logic* 2(4): 418–446.
- Lewis, D. (1979). 'Counterfactual dependence and time's arrow', *Noûs* 13(4): 455–476.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Wiley-Blackwell.
- Mendelson, E. (1987). *Introduction to Mathematical Logic*. Monterey, CA: Wadsworth & Brooks.
- Mill, J. (1967). *A System of Logic*. London: Longmans.
- Plantinga, A. (1974). *The Nature of Necessity*. Oxford: Clarendon Press.
- Ramsey, F. (1929). 'Law and causality', in his *Philosophical Papers*, ed. D. H. Mellor. Cambridge: Cambridge University Press 1990, pp. 140–163.
- Rosen, G. (1990). 'Modal fictionalism', *Mind* 99(395): 327–354.
- Stalnaker, R. (1968). 'A theory of conditionals'. In N. Rescher (ed.), *Studies in Logical Theory* (American Philosophical Quarterly Monographs 2). Oxford: Blackwell, pp. 98–112.
- Stalnaker, R. (1984). *Inquiry*. Cambridge: Cambridge University Press.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.

A Commentary on ‘Necessity, Conditionals and Apriority’

Dorothy Edgington

1. Introduction

Keith came to Birkbeck early in the 1980s to enquire about resuming studies in philosophy after a long gap. He actually applied to do an MA. He submitted some impressive work on the status of Euclidean geometry. After discussion it was decided that the best plan was for him to see if he could re-register for his long-abandoned Oxford doctorate, which is what happened. As he lived and worked in London, I was appointed his supervisor, and he was, *de facto*, an active Birkbeck graduate student. I learned more from him than he did from me, but we got on very well, and the doctorate was completed. I was pleased when he got a job in the philosophy department at King’s College London, and even more pleased when, some years later, he returned to a job at Birkbeck, to which he had a strong attachment. He is a terrific person to talk philosophy with.

Keith’s philosophy is always ambitious, original and bold. His chapter in this volume is no exception. Its aim is to prove that the necessary and the *a priori* coincide, and thus to ‘reduce the modal to the epistemic’ (p. 125). Its method is as follows. Necessity is defined in terms of unconditionality: if *C* is necessary, then whatever else were the case, *C* would be the case; *C* is true ‘come what may’; for every antecedent *A* ‘If it were that *A*, it would be the case that *C*’ is true. Truth conditions for counterfactual conditionals are to be given not in terms of possible worlds (for this would be to take modal notions for granted), but in terms of the consequent’s being inferrible from the antecedent, together, in general, with additional premises. Nelson Goodman’s (1955) attempt to give truth conditions for counterfactuals along these lines is the model, supplemented in a new way.

Keith endorses an inference-based account of conditionals in general. I shall make some preliminary comments about this. First, if we infer *C* from *A* together with some background premises *S*, then for this to suffice for the truth of ‘If *A*, *C*’, our background premises had better be true; there is no guarantee that the conditional is true if *C* is inferred from *A* together with false premises.

Second, if we want the inferribility-conditional to be stronger than the material conditional, we need to be selective as to which truths it is permissible to use as

background assumptions. Suppose any truth will do. Then the inferribility-conditional is no stronger than the material conditional. For whenever $\neg A \vee C$, i.e. $A \rightarrow C$ is true, C is inferrible from A .

Third, Keith moves immediately from inferribility to entailment: “If A , then C ” is true if the antecedent A , possibly together with further relevant conditions S , ... entails the consequent C (p. 117). But many inferences are not entailments – inference to the best explanation, for example. For instance ‘They’re not at home; because the lights are off; and *if they had been at home the lights would have been on*’. There is no entailment between antecedent and consequent. Suitable inferences which are entailments may give sufficient conditions for the truth of the conditional, but there are plenty conditionals which we believe – especially those concerning human behaviour – such that there is no entailment of consequent from antecedent together with other true premises.

2. Goodman

Keith follows Goodman in giving an entailment-based theory of counterfactuals. According to Goodman, ‘If A were the case, C would be the case’ is true if and only if there is a suitable set of true premises S such that $A, S \vdash C$.

Consider Goodman’s stock example,

- (1) If the match had been scratched, it would have lit.

If the match is actually dry at the relevant time, oxygen is present, there is no gale blowing, the match is well made, etc. and there is a law of nature connecting these conditions with the match’s lighting on the counterfactual supposition that it was scratched, the counterfactual is true.

Introducing the counterfactual case, Keith qualifies the ‘relevant circumstances’ with an enigmatic parenthetical ‘now hypothetical’ (p. 118). Goodman could not be clearer that the conditions are not hypothetical, they have to be true: ‘We do not assert that the counterfactual obtains *if* the circumstances obtain; rather, in asserting the conditional we commit ourselves to the actual truth of the statements describing the requisite relevant conditions’ (1955: 8).

When you make a counterfactual supposition, e.g. that the match was scratched, clearly you need to give up some truths – such that the match remained motionless and untouched on the table. But, as we saw, you need to rely on other truths, such as that the match was dry. Which truths can one rely on, when one makes a counterfactual supposition? Using Goodman’s excellent term, which truths are *cotenable* with the counterfactual supposition? Some cases are clear: any truth inconsistent with your antecedent is not cotenable with the antecedent. Another case handled by Goodman, and cited by Keith, arises because of examples like the following. When Ann goes to her office, she sometimes takes the lift, and sometimes climbs the stairs, with no fixed pattern. She did not go to her office yesterday. She did not take the stairs yesterday, and she did not take the lift. Suppose she had gone to her office yesterday. You can’t hold on to the truth that she did not take the stairs, and infer that she would have taken the lift,

nor can you hold on to the truth that she did not take the lift, and infer that she would have taken the stairs. For, had she gone to her office, she might have done either. So, says Goodman, if you can derive C from A together with S , and $\neg C$ from A and some other S , neither S nor S is cotenable with A .

But Goodman ran into trouble in trying to give a general answer to the question of cotenability. (1) above is true, because the match was dry, there was oxygen, etc., and there is a law of nature such that, together with these facts, the supposition that it was scratched entails that it lit. Now consider

(2) If the match had been scratched, it would not have been dry.

(2) is false. But the consequent follows from the antecedent, a law of nature and other facts including the fact that *the match did not light*.

The difference between the two cases must lie in cotenability. A truth S is cotenable with a supposition A if it's not the case that S would not be true if A were true (Goodman 1955: 13). That the match was dry is cotenable with the supposition that it was scratched, because it is not the case that if the match had been scratched, it would not have been dry. That the match did not light, though true, is not cotenable with the supposition that it was scratched, because it is the case that if the match had been scratched, it would have lit. Why is (1) true and (2) false? Because (2) is false and (1) is true. We have a tight, unexplanatory circle.

Keith's idea is to import into Goodman's account of counterfactuals, adapt and rename, a central notion from the rival possible-worlds accounts of Robert Stalnaker (1968) and David Lewis (1973) – that of a relation of comparative similarity, or closeness to actuality, of possible worlds. It is Lewis's semantics that he follows, so let us turn briefly to that.

3. Lewis

Lewis's monograph begins thus:

'If kangaroos had no tails they would topple over' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over. I shall give a general analysis of counterfactual conditionals along these lines. (1973: 1)

And a little later:

We might think it best to confine our attention to worlds where kangaroos have no tails and *everything else* is as it actually is; but there are no such worlds. Are we to suppose that kangaroos have no tails but their tracks in the sand are as they actually are? Then we shall have to suppose that these tracks are produced in a way quite different from the actual way. Are we to suppose that kangaroos have no tails but that

their genetic makeup is as it actually is? Then we shall have to suppose that genes control growth in a way quite different from the actual way (or else there is something, unlike anything there actually is, that removes the tails). And so it goes; respects of similarity and difference trade off. If we try too hard for exact similarity to the actual world in one respect, we will get excessive differences in some other respect. (ibid.: 9)

Lewis's truth conditions are as follows. A counterfactual "If it were that *A*, it would be that *C*" is vacuously true if *A* is true at no possible world. It is non-vacuously true if and only if some world in which 'both *A* and *C* are true is more similar to our actual world, overall, than is any world where *A* is true but *C* is false' (Lewis 1979: 41).

Later in the chapter, I sometimes express Lewis's second condition thus: 'If it were that *A*, it would be that *C*' is non-vacuously true if and only if *C* is true at all the closest *A*-worlds (and I apply this to Keith's adaption of Lewis). My formulation is almost always equivalent to Lewis's formulation above, but, he tells us, not quite always, because there can be infinitely descending chains of closer and closer *A*-worlds, and so no closest *A*-world. His example (Lewis 1973: 20–21): a counterfactual which begins 'If this line were over an inch long, ...', when actually it is a little under an inch long. There is no *closest* world in which it is over an inch long. No harm is done by my setting aside these cases.

In his early work, Lewis seemed to be appealing to similarity in a common-or-garden sense – vague and sometimes indeterminate, no doubt, but so are counterfactuals. He used as analogies similarities between cities, and between faces. Counterexamples arose, the most famous being Kit Fine's (1975) 'If Nixon had pressed the button, there would have been a nuclear holocaust'. Any counterfactual of the form 'If such-and-such had happened, things would have been very different' causes problems. Lewis (1979) responded by giving a much more refined account of the respects of similarity that matter, on what he called the 'standard resolution of vagueness' of counterfactuals. (He did not claim that this was the only way of assessing counterfactuals, but that it was the standard, default way.) For typical counterfactuals his refined account has these consequences: the relevant worlds to consider – the closest *A*-worlds – are those exactly like the actual world until shortly before the antecedent time, when, if necessary, a small violation of the actual laws is needed to get the antecedent true. The actual laws are sacrosanct after that. Once we have departed from perfect match with the actual world, 'it is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly' (1979: 47–48).

This gets round the Nixon example and others like it. It also brings Lewis closer to Goodman. It is stated in terms of what you hold constant, and also gives laws of nature a more prominent role.

4. Far-fetchedness

Similarity between possible states of affairs, or possible worlds, their closeness to the actual world, is all very well, but it doesn't sound right for a relation between propositions, so Keith speaks instead of 'far-fetchedness' (the more similar to the actual world, the less far-fetched, and vice versa). What is to be far-fetched? My dictionary

(Chambers) says ‘forced or unnatural, unlikely, improbable, unconvincing’. At first sight it is ‘unlikely, improbable, unconvincing’ that seem most promising: when I think of clear cases of the far-fetched, they are all improbable, and in combination with other improbable things, become more improbable, as in Keith’s example of the talking donkey (p. 120). It is certainly a valuable exercise to determine what are the more likely, which the less likely, outcomes of a counterfactual supposition. But this does not yield a truth condition. The more likely outcome isn’t always the one that happens, and in the counterfactual case, the more likely outcome might not be the one that would have happened. Jane is told on authority that it’s 90 per cent likely that she will be cured if she has the operation. She declines the operation. It’s 90 per cent likely that she would have been cured if she had had the operation. That doesn’t make it true that she would have been cured, for she might not have been. Similarly, in a mini-lottery, I hold three tickets, two other people hold one each. An emergency arises so the lottery is called off. The most likely outcome is that I would have won, had the lottery gone ahead, but we cannot say that it is true that I would have won. So, if it is to yield truth conditions, far-fetchedness is not to be understood in terms of improbability. Keith told me he agrees, and that is not how it is to be understood.

Keith could go down Lewis’s road in specifying criteria for far-fetchedness. But he chooses not to. We do not need, and should not expect, a reductive analysis of the counterfactual conditional, he claims. So he defines far-fetchedness in terms of the counterfactual, thus: *P* is less far-fetched than *Q* iff, if exactly one of *P*, *Q* were true, *P* would be true and *Q* would be false (p. 120). Apply this to our problem with the match: suppose it had been scratched; why should one hold on to the fact that the match was dry, and conclude that it would have lit, rather than hold on to the fact that it didn’t light, and conclude that it wouldn’t have been dry? Because ‘scratched, dry, lit’ is less far-fetched than ‘scratched, not dry, didn’t light’. What does ‘far-fetched’ mean here? It means that if exactly one of the above triples were true, it would be the former and not the latter.

This is still circular and non-explanatory; and one wonders whether the detour through far-fetchedness adds anything significant to Goodman’s circular account. Well, having adopted the ordering structure of Lewis’s account, this does have implications for the logic of counterfactuals. Famously, it engenders counterexamples to transitivity, strengthening of the antecedent, and contraposition (see Lewis 1973: 31–36, Stalnaker 1968: 38–39). For instance, we can have all closest *A*-worlds are *C*-worlds; but not all closest *A*&*B*-worlds are *C*-worlds. As all *A*&*B*-worlds are *A*-worlds, we have a counterexample to transitivity.

I think that Keith’s definition of far-fetchedness is not quite correct, for the following reason. I should have thought that all true propositions have minimal, or zero, far-fetchedness: you don’t have to depart from actuality at all for them to be true. (Compare: the actual world is more similar to the actual world than any other world is.) But, for two true propositions, it is often clear which one would be true, if exactly one of them were true. Let *O* be, there is oxygen near the Earth’s surface. Let *B* be, I ate a banana at lunch. Both are true, and if exactly one were true, it would be *O*, not *B*. So, on Keith’s definition, *O* is less far-fetched than *B*, so not all truths are minimally far-fetched. Thus, I think his definition of far-fetchedness should be restricted to false propositions.

Indeed, if, as he later claims, his truth conditions coincide with Lewis's, that restriction is necessary, because Lewis's theory has the principle known as 'centering' (1973: 14–15): 'If it were that *A*, it would be that *C*' is always true when *A* and *C* are both true; and this is not obviously the case if truths can be more or less far-fetched.

5. 'Relevance'

Keith gives a definition of 'relevance' which plays a large part in the rest of his chapter, and the proofs in the Appendix:

A proposition *S* is said to be *relevant to A* if *S* and *A* are consistent, and every proposition consistent with *S* is less far-fetched than any proposition inconsistent with *S*. (p. 119)

This definition puzzled me for some time. It didn't seem right. Let *A* be the proposition that the match was scratched. One would think that the proposition that the match was dry was relevant, or part of what was relevant, to *A*. So the proposition that the match was wet is inconsistent with *S*. But I can think of many propositions, consistent with *S*, that are *more* far-fetched than that the match was wet: that I ate fried cockroaches for breakfast, that I ran a mile in three minutes, that the moon is made of green cheese. . . . So it's not the case that every proposition consistent with *S* is less far-fetched than any proposition inconsistent with *S*. We could avoid this only if we included the negations of these far-fetched propositions – that I did not eat cockroaches for breakfast, etc., in *S*.

But then I realized that they would have to be there in *S* anyway. Goodman was investigating which truths one can rely on to get from a particular *A* to a particular *C*. What is relevant to getting from *A* to *C* may not be relevant to getting from *A* to *D*. What is relevant to 'If you had scratched the match, it would have lit' differs from what is relevant to 'If you had scratched the match, you would have been punished' (whether or not it lit; you have been told not to play with matches). Keith, by contrast, is defining what is 'relevant to *A*'; full stop. *S* is an all-purpose condition that can be appealed to for any consequent whatsoever. The conditional to be assessed might be 'If you had scratched the match, I would have eaten cockroaches for breakfast'. Keith's all-purpose *S* must be at least a maximally complete least-far-fetched proposition which would be true if *A*. But normally (I think always) there is more than one such least-far-fetched maximal proposition, because *A* can come about in more than one way. There are different ways Ann could have gone to her office, Jane's operation might have gone, or even, the match could have been scratched. So *S* must be a disjunction of all least-far-fetched propositions describing how things would be if *A* were true. All this machinery does come later in Keith's chapter (p. 121), when arguing that his theory and Lewis's coincide. He calls these maximal consistent propositions 'histories'.

A point about Keith's 'histories': are they not, pretty much, what most philosophers mean by possible worlds? All but one is not an actual history, but a possible history. Lewis's theory of counterfactuals is popular, but his theory of the nature of worlds

(Lewis 1986) is not. For Lewis, non-actual possible worlds are just as real as the actual world, *viz.* 'I and all my surroundings', but spatio-temporally unconnected to it. Few philosophers accept this. For instance, Jonathan Bennett (2003), introducing his account of counterfactuals, having said that he cannot accept Lewis's extreme modal realism, says that according to the abstract modal realism he adopts, 'a possible world is a *consistent* proposition *P* which is *maximal*, meaning that every proposition is either entailed by *P* or inconsistent with it. Or, in yet other words, it is a *total way things might be*' (156). On this view, 'the actual world' is not 'I and all my surroundings', but everything that is the case – the true maximal consistent proposition. Bennett also says 'I shall regularly write as though the worlds in question *had* histories, whereas strictly speaking they *are* histories' (*ibid.*: 158). And he goes on to develop a broadly Lewisian account of counterfactuals.

I wish to point out that the structure of the inference from *A* to *C* via *S* on Keith's account is very different from its structure in Goodman's theory. For Goodman, we try to find suitable *truths* – facts and laws – which, together with the assumption of *A*, entail *C*. 'If *A* were the case, *C* would be the case' is true *because of* facts, *P*₁, *P*₂, etc. On Keith's theory, *S* is false. It represents what *would be true if A*. It is a disjunction of all the least-far-fetched histories in which *A* is true. And as Lewis said, if *A* is actually false, there are no worlds/histories that differ in *only* that respect. All the least-far-fetched histories in which *A* is true differ from the actual history in many respects. (Recall Lewis's remarks about the kangaroos, and also his claim that if *A* was determined to be false by the past and laws, we need a 'small miracle' – violation of the actual laws – in all the closest *A*-worlds.) Similarly for Keith's histories: each disjunct in the disjunction of least-far-fetched histories contains falsehoods.

Unlike Goodman's, the structure of Keith's argument from *A* to *C* via *S* is this: suppose *A* were true; *then as a consequence* (of a particular philosophical theory) *S*; i.e. one of these least-far-fetched histories *would be true* (call this inference *I*); and *C* is true in all of them; therefore *C*. *S* is no longer a premise, though it itself rests on the unstated premise of a philosophical theory involving far-fetchedness. We shall return to inference *I* when we look at Keith's account of logical consequence.

6. Relative and absolute correctness

I accept Keith's argument that Lewis's theory and his revised version of Goodman's theory are equivalent (§3, and Appendix, Theorem 2): the same counterfactuals come out true on both. This is not very surprising, given the way 'relevance' is defined, and the one-one correlations between worlds and histories, closeness and far-fetchedness. I am not sure that this construction of Keith's deserves to have Goodman's name attached to it.

However, I do not accept his proof that (his version of) Goodman's theory is correct (§6, and Appendix, Theorem 3). The proof is, as he fully recognizes, circular, as it only speaks to those who already accept that an ordering by closeness/far-fetchedness is a key to the interpretation of counterfactuals. The proof uses what he calls 'known principles of the logic of counterfactuals' (p. 123). These axioms involve the ordering of

propositions by far-fetchedness (p. 123). But not all theories of counterfactuals employ an ordering relation on worlds/propositions. Stalnaker's and Lewis's work have been influential and popular, but they do not exhaust the field. There are those who construe the counterfactual as a strict conditional over a contextually determined set of worlds. They include William Lycan (2001), Anthony Gillies (2007) and, recently, Timothy Williamson (2020). They thus reinstate the validity of strengthening the antecedent, transitivity and contraposition: the contextually determined set of worlds is kept constant throughout an inference; apparent counterexamples involve a shift in context. They also disagree with Lewis on 'centering': on these theories, the truth of *A* and of *C* is not sufficient for the truth of 'If it were that *A*, it would be that *C*'.

Second, there are those who theorize in terms of a probability distribution over the possible outcomes of a counterfactual supposition, rather than an ordering of them. They include Ernest Adams (1975), Brian Skyrms (1980), Hannes Leitgeb (2012), Richard Bradley (2017) and myself (Edgington 2008). As I put it once, why put all your eggs in the closest baskets? (1995: 259). A complaint against both Lewis's and the strict-conditional account is that they go plain false too easily. I may be highly confident that if you had tossed the coin ten times you would have got at least one head, or that Jane would have recovered if she had had the operation, but as the consequent is not true in *all* the closest antecedent-worlds, on these accounts these counterfactuals are definitely false.

Third, there is a 'premise semantics' for counterfactuals developed by Angelika Kratzer (1981), popular with linguists, which does not involve orderings, although Lewis (1981) showed that it is equivalent to an ordering semantics.

7. The necessary and the *a priori*

Keith follows Mill in explaining necessity as *unconditionalness*: the necessary is what is true, whatever supposition we make regarding all other things. The necessary is counterfactually robust: if *C* is necessary, then for any *A* at all, if *A* were true, *C* would be true. What about the case in which *A* is $\neg C$. Then, (even) if $\neg C$ were true, *C* would be true. (Keith takes this formulation from Williamson 2007: 151–56.) This has a paradoxical ring to it – a hint that Mill's insight is being pushed too far – but I do not want to deny that it is true.

A proposition is *a priori*, he tells us, if it is primitively evident, or a logical consequence of primitively evident truths. *C* is a logical consequence of a set of premises Σ if there is a proof of *C* from Σ . A proof is a chain of propositions each of which is either primitively evident, or one of Σ , or stands to preceding propositions in a relation that is evidently truth-preserving. He says 'I will write " $\Sigma \vdash C$ " to mean that there is a proof of *C* from Σ . If *C* can be proved without the need for any premises I write " $\vdash C$ ". ... A proposition is said to be *a priori* if $\vdash C$ ' (pp. 125–26). This is an epistemic notion, not a modal notion, he tells us.

The symbol ' $\vdash$ ' occurs from the beginning of Keith's chapter and throughout, and is initially introduced as entailment, so I assume that '*A*, *S* entail *C*' means the same as '*C* is a logical consequence of *A*, *S*, in the sense that the latter is elucidated here.

Go back to inference *I* in Keith's account of the counterfactual (see above, end of my §5): suppose it were the case that *A*; then as a consequence, *S*, i.e. one of these least-far-fetched histories would be true (and *S* entails *C*). *S* is not a logical consequence of *A*, as Keith defines it. It rests on a substantive philosophical theory, one of considerable merit, but it cannot be said to be primitively evident that it is true. It is not primitively evident that an ordering of propositions by far-fetchedness is the key to counterfactuals. No substantive philosophical theory is primitively evident. Lewis did not think his theory of counterfactuals in terms of close possible worlds was primitively evident. In the Introduction to Volume I of his *Philosophical Papers*, some remarks on philosophical method begin:

The reader in search of knock-down arguments in favor of my theories will go away disappointed. Whether or not it would be nice to knock down disagreeing philosophers by sheer force of argument, it cannot be done (1983: x).

It is hard to have much confidence in this notion of the primitively evident. Almost every principle of logic is disputed by someone – indeed, someone who is highly intelligent and perfectly sincere. There is intuitionist logic, relevance logic, quantum logic, dialethic logic, three-valued logics (e.g. for vagueness and indeterminacy) and so on. What is primitively evident to a classical logician may not be primitively evident to, for example, Michael Dummett (1977), Alan Ross Anderson and Nuel Belnap (1975), or Graham Priest (2006), who make serious cases for alternatives. In arguing against alternatives, one is arguing that classical logic does a better overall job of systematization. But we have left the 'primitively evident' behind. Another reason for caution is that there are things which strike everyone as primitively evident, such as Frege's axiom 5, that for every predicate there is a set of things which satisfy it, until they are proved to be false. Euclid's Axiom of Parallels is another example. Even the claim that counterfactual conditionals are transitive had a 'primitively evident' feel before the work of Adams, Stalnaker and Lewis cast it into doubt.

8. The dénouement: the necessary and the a priori coincide

This is remarkably brief. Suppose *A* is a priori. Then $\vdash A$. Then $\neg A \vdash A$. So $\neg A \Box \rightarrow A$. So *A* is necessary.

Conversely, suppose *A* is necessary. Then $\neg A \Box \rightarrow A$. So either $\neg A \vdash A$, or there is a relevant *S* such that $\neg A, S \vdash A$. But if the first disjunct does not hold, the second cannot hold either, for then *S* would be inconsistent with $\neg A$, and so not cotenable with $\neg A$. Therefore, $\neg A \vdash A$. So $\vdash A$. So *A* is a priori (§9).

My initial reaction was, how can this be? Surely, every statement of pure number theory is either necessarily true or necessarily false. Yet there is no guarantee that every true universally quantified statement of pure number theory is provable. There can be properties that hold of every natural number, without there being a proof that this is so. Take Goldbach's Conjecture (*G*), as yet unproven: every even number greater than 2 is the sum of two prime numbers. Suppose it is true. Then it is necessarily true. It is

counterfactually robust: whatever else were the case, it would be true. Then, according to Keith, it is provable from self-evident premises. But perhaps there is no way of proving it to be true.

Full second-order logic is another example. It is incomplete in the sense that there are logically valid statements which are not provable.

Not everyone uses entailment in Keith's sense. Another elucidation of entailment is necessary preservation of truth. This leaves open the question whether all necessary truths are a priori. Suppose G is true. Then it is necessarily true – it is impossible that it is false. Hence no argument with G as conclusion has true premises and a false conclusion: for any A , $A \vdash G$ and $\neg A \vdash G$. So $\vdash G$, meaning that G is necessary. We could go through Keith's steps: suppose G is true; so $\neg G \Box \rightarrow G$; so $\neg G \vdash G$; so $\vdash G$. But that would not mean that G is knowable a priori, but rather, that it cannot be false.

However, in the final version of his chapter, Keith allows that proofs can be of infinite length, in which case Goldbach's Conjecture is not a counterexample to the claim that all necessary truths are a priori. He refers the reader to his book (Hossack 2020: 31–32), where he advocates the rule of Complete Induction, 'If for each element x of the domain you have proved Fx , you may infer $\forall x Fx$ ' – even if no finite being could employ the rule as a means of coming to know the conclusion of a proof. Suppose Goldbach's Conjecture is true. Then infinitely many steps of checking the even numbers greater than 2 will show that each is the sum of two prime numbers. So Goldbach's Conjecture is provable, hence a priori, after all.

I don't care for the idea that proofs can be of infinite length. Proof is an epistemic notion, and should be a procedure by which one can, at least in principle, come to know its conclusion. In any case, while Goldbach's Conjecture is a nice example because of its remarkable simplicity, mathematics is full of unsolved, perhaps insoluble, problems, and it is not clear that the rule of Complete Induction would be sufficient to solve them all. The question whether there are always smooth solutions to the Navier-Stokes equations – a system of nonlinear partial differential equations, arising in fluid dynamics – may be such an example.

Leaving mathematics aside, there is the case of identity statements. These do seem to be counterfactually robust: if x is y , then anything that would have been true of x in a counterfactual circumstance, would have been true of y in that counterfactual circumstance. Hence, there is no counterfactual circumstance in which x is not y . But it is not, in general, knowable a priori that x is y .

What about the other direction? Here is something which seems as trivially self-evident as anything, hence satisfying Keith's notion of the a priori: what Dee weighs at t , is her actual weight at t . But this is not counterfactually robust: there are counterfactual circumstances in which what Dee weighs at t would be less than her actual weight at t . If she had stuck to her diet for another month, she would have weighed less than what she actually weighs. If she had been significantly even more self-indulgent, she would have weighed more than what she actually weighs. Hence, though a priori, it is not necessary that Dee's weight at t is her actual weight at t .

Note that, although the above statement is not counterfactually robust, it is what one might call factually robust: whatever else *is* the case, if Dee weighs m at t , her actual weight at t is m . This has led some philosophers to postulate two notions of necessity,

one epistemic, the other not invariably so. See Martin Davies and Lloyd Humberstone (1980). I find this view attractive, but there is still the difficulty I mentioned above – that the notion of the primitively evident seems insecure.

There is another, more austere, conception of logical truth and logical consequence, which makes no mention of the necessary *or* the a priori, but simply rests on universality. A logical truth is a proposition of a form such that all propositions of that form are true. A logically valid pattern of inference is a formal pattern of inference none of whose instances leads from true premises to a false conclusion. I think this was the notion to which Goodman appealed. While he does not assimilate causal laws and logical laws, there is no indication that he thought of them as radically different in status. He writes of his relevant conditions ‘leading by law’ from A to C, meaning by causal law and/or by logical law (1955: 13). To call Goodman ‘Quinean’ is true but a little misleading: it is said that he influenced Quine’s thinking about these matters, in ways which led to Quine’s ‘Two Dogmas of Empiricism’ (1951). See Cohnitz and Rossberg (2020).

In sum, while I have great admiration for the ingenuity of Keith’s chapter, and the skill with which it is developed, I don’t feel forced to the conclusion that the necessary and the a priori coincide. As Lewis said, whether or not it would be nice to knock down disagreeing philosophers by sheer force of argument, it cannot be done.

References

- Adams, E. W. (1975). *The Logic of Conditionals*. Dordrecht: Reidel.
- Anderson, A. R. and Belnap, N. D. (1975). *Entailment: The Logic of Relevance and Necessity*. Princeton: Princeton University Press.
- Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.
- Bradley, R. (2017). *Decision Theory with a Human Face*. Cambridge: Cambridge University Press.
- Cohnitz, D. and Rossberg, M. (2020). ‘Nelson Goodman’, *The Stanford Encyclopedia of Philosophy*, Edward Zalta (ed.), <https://plato.stanford.edu/entries/goodman/>
- Davies, M. and Humberstone, L. (1980). ‘Two notions of necessity’. *Philosophical Studies* 38: 1–30.
- Dummett, M. (1977). *Elements of Intuitionism*. Oxford: Oxford University Press.
- Edgington, D. (1995). ‘On conditionals’. *Mind* 104: 235–329.
- Edgington, D. (2008). ‘Counterfactuals’. *Proceedings of the Aristotelian Society* 108: 1–21.
- Fine, K. (1975). Critical notice of David Lewis’s *Counterfactuals*. *Mind* 84: 451–458.
- Gillies, A. S. (2007). ‘Counterfactual scorekeeping’. *Linguistics and Philosophy* 30(3): 329–360.
- Goodman, N. (1955). *Fact, Fiction and Forecast*. Indianapolis: Bobbs-Merrill.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number*. London: Bloomsbury.
- Kratzer, A. (1981). ‘Partition and revision: The semantics of counterfactuals’. *Journal of Philosophical Logic* 10: 201–216, reprinted in Kratzer, A. (2012). *Modals and Conditionals*, Oxford: Oxford University Press, pp. 72–84.
- Leitgeb, H. (2012). ‘A probabilistic semantics for counterfactuals, part A’. *Review of Symbolic Logic* 5: 16–84.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.

- Lewis, D. (1979). 'Counterfactual dependence and time's arrow'. *Noûs* 13: 455–476, reprinted in Lewis, D. (1986). *Philosophical Papers* Volume 2. Oxford: Oxford University Press, pp. 32–52. Page references to the reprint.
- Lewis, D. (1981). 'Ordering semantics and premise semantics for counterfactuals'. *The Journal of Philosophical Logic* 10: 217–234, reprinted in Lewis, D. (1998). *Papers in Philosophical Logic*. Cambridge: Cambridge University Press, pp. 77–110.
- Lewis, D. (1983). *Philosophical Papers* Volume 1. Oxford: Oxford University Press.
- Lewis, D. (1986). *The Plurality of Worlds*. Oxford: Oxford University Press.
- Lycan, W. (2001). *Real Conditionals*. Oxford: Oxford University Press.
- Priest, G. (2006). *In Contradiction: A Study of the Transconsistent*, 2nd edn. Oxford: Oxford University Press.
- Quine, W. V. O. (1951). 'Two dogmas of empiricism'. *Philosophical Review* 60: 20–43, reprinted in Quine, W. V. O. (1961). *From a Logical Point of View*. New York: Harper and Row, pp. 20–46.
- Skyrms, B. (1980). *Causal Necessity*. New Haven, CT: Yale University Press.
- Stalnaker, R. (1968). 'A theory of conditionals', in *Studies in Logical Theory, American Philosophical Quarterly Monograph Series*, 2. Oxford: Basil Blackwell, pp. 98–112, reprinted in Jackson, F. (ed.) (1991). *Conditionals*. Oxford: Clarendon Press, pp. 28–45. Page references to the reprint.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.
- Williamson, T. (2020). *Suppose and Tell*. Oxford: Oxford University Press.

The Mathematicians' Use of Diagrams in Plato¹

Tamsin de Waal

1. Introduction

I was lucky enough to have Keith as one of my supervisors, whilst a doctoral student at King's College London. My thesis was primarily focused on Plato and the problem of how we can come to know forms, but related to the access problem in contemporary philosophy of mathematics. Keith introduced me to a world that I was previously unfamiliar with, with infectious enthusiasm. Our discussions were always wide-ranging, but, somehow, we always seemed to come back to Euclid. So it seemed only appropriate that my chapter in this volume should be about diagrams, and the role they play in coming to know forms.

It is clear that mathematics, on Plato's view, has an important role to play in facilitating progress towards knowledge of forms. This is particularly clear in the *Republic*. Here, mathematics constitutes a significant part of the prospective philosopher-rulers' education programme – both in the sense that it takes up a lot of space in the description of the programme, and in the sense that it takes up so many (ten) years of training. And it is proposed on the grounds that such studies have the power to effect 'a conversion and turning about of the soul' (521c), and to 'draw the soul away from the world of becoming to the world of being' (521d).

This raises a number of broad questions. Why should Plato set such mathematical store on mathematics? How does he see it effecting cognitive development? How much of a cognitive shift is it capable of effecting? How do the different mathematical disciplines contribute to this shift? Is mathematics the only thing that can prompt this cognitive development? Whilst I engage with these broad questions here, the focus of the chapter is a relatively narrow one: whether, on Plato's account, the use of diagrams in mathematics plays a part in facilitating cognitive development.

Plato highlights the use of diagrams in mathematics in the context of the *Republic* divided line image (509dff), itself, of course, a diagram. Following immediately on from the sun simile (506eff), the divided line picks up on the distinction drawn there between a visible and an intelligible realm: 'Well, suppose you have a line divided into two unequal parts, and then divide the two parts again in the same ratio, to represent the visible and the intelligible realms' (509d).² The four sub-divisions correspond to four distinct cognitive states – intelligence (*noēsis*) and understanding (*dianoia*), which

are correlated to the intelligible realm, and belief (*pistis*) and imagination (*eikasias*), which are correlated to the visible realm (511d–e).³ These are to be arranged ‘in a proportion, considering that they participate in clearness and precision in the same degree as their objects partake of truth and reality’ (511e). It is generally assumed that we should picture a vertical line, here,⁴ and that the larger part is the upper part, and represents the intelligible realm. But it is notable that none of this is explicitly specified.

Dianoia is distinguished from *noēsis* on two methodological grounds:

- Whilst those in a state of *noēsis* treat hypotheses as just that and seek ‘that which requires no assumption and is the starting-point of all’ (511b), those in a state of *dianoia* make use of hypotheses, using them as first principles and deeming them unnecessary to explain (510c).⁵
- Further, whilst those in a state of *noēsis* rely on forms only and progress systematically through these, those in a state of *dianoia* make use of sensible objects – they investigate by using these (the originals of the visible realm) as images (510b).

In order to clarify the distinction between *dianoia* and *noēsis*, Socrates identifies the latter with dialectic, and the former with mathematics, the sort of reasoning that mathematicians engage in.⁶ Mathematicians, ‘students of geometry and reckoning and such subjects’ (510c), rely on hypotheses, and:

They use the visible forms (*tois horōmenois eidesi*) besides, and make their accounts about them, not thinking about/intending them, but those things to which these are like (*eoike*), making their accounts for the sake of the square itself and the diagonal itself, not for the sake of this (diagonal) which they draw (*graphousin*), and likewise in regard to the other things: these very things which they mould and draw, of which there are shadows and images (*eikones*) in water, they are using in their turn as images (*hōs eikosin*), while they are seeking to get a view of the things which one can see in no other way than with thought.

510d–e

The references to drawing, and, more particularly, to the drawing of squares and diagonals, make it clear that ‘the visible forms’, the sensible objects, in question are, first and foremost, geometrical diagrams.⁷ Mathematicians use diagrams and treat these sensible objects as images or likenesses. The reference to ‘students of geometry *and reckoning and such subjects*’ makes it clear that the use of geometrical diagrams is not confined to geometry.

The use of diagrams here has received very little attention. Commentators have tended to pass over this, focusing instead on the use of hypotheses.⁸ In as much as the use of diagrams is noted, it is generally cast in a predominantly negative light – as either a defect of mathematical method, or a failure of certain mathematicians.⁹ Plato is clear, it is true, that knowledge of forms can only be achieved where sensible objects and hypotheses are relinquished, so, in this sense, the use of diagrams in mathematics is cognitively limiting. However, as I aim to demonstrate in this chapter, the use of

diagrams in mathematics also has a highly positive cognitive effect, on Plato's account, and plays a crucial role in facilitating progress towards knowledge of forms.

2. What mathematics is Plato recommending?

The first point to consider is whether the mathematical method described in the divided line, complete with the use of sensible objects, is the same mathematics recommended for the prospective philosopher-rulers, the guardians, in Bk VII of the *Republic*. To put this another way, should we identify the dianoetic method of the divided line image with the mathematical method, correctly applied?

(i) The dianoetic versus mathematical method

The mathematical curriculum recommended for the guardians begins with arithmetic and logistic.¹⁰ This is to be followed by plane geometry (526c–527c), three-dimensional (solid) geometry (528a–d), astronomy (527d–528a) and harmonics (530–531c). Socrates pointedly insists that the five mathematical subjects be studied in this order.¹¹

In setting out the mathematical curriculum, Socrates distinguishes the mathematics to be studied by the guardians both from applied mathematics (the sort of mathematics generals might use – i.e. what the *Philebus* classifies as popular, as opposed to philosophical, mathematics),¹² and from mathematics as practised by contemporary mathematicians. With respect to the latter, Socrates picks out astronomy (528eff) and harmonics (531aff) for particular criticism – these concern themselves with the visible and the audible, respectively, proceeding by means of the senses, rather than by reasoning. Solid geometry, he observes, is altogether neglected by contemporary mathematicians (528b–d). The mathematical method recommended (i.e. the philosophical mathematical method) is one that proceeds by means of problems (530b–c),¹³ concerning itself with intelligible entities,¹⁴ and one that is pursued for the sake of knowledge (525c–d).

As to whether we should identify this method with the dianoetic method, on one possible reading, the dianoetic method is a *mis*application of the mathematical method by certain mathematicians – for example, the mathematicians of the day.

On this reading, the correctly applied mathematical method does not take its hypotheses to be first principles and does not make use of sensible objects. But in this case, what *is* mathematics, as distinct from dialectic? On a more moderate version of this reading, the correctly applied mathematical method does make use of sensible objects, but does not, like the dianoetic method, make *inappropriate* use of these.¹⁵ It is not obvious, though, what might constitute 'inappropriate' use of sensible objects. The most significant sense in which sensible objects might be deemed to be misused is where they are treated not as the likenesses they are, but as originals. This is what the astronomers of the day are accused of doing at 528eff. Plato, though, is clear that the dianoetic mathematicians treat sensible objects as likenesses.¹⁶ And it is notable that Plato does not criticize the geometers (or arithmeticians) of the day in the way he criticizes the astronomers of the day, and those who practise harmonics.

Alternatively, then, we might take it that the dianoetic method *is* the mathematical method, correctly applied – after all, the philosophical mathematician of *Republic* Bk VII, is, as we shall see, characterized as being in a state of *dianoia*.

On this reading, the use of hypotheses and diagrams are simply necessary features of the mathematical method. And in this case, Plato cannot be criticizing (any) mathematicians, for relying on these. Might he nonetheless mean to be critical of these features? It is hard to see why he would so strongly advocate the study of mathematics if he meant to criticize features that are fundamental to the practice of mathematics. On this reading, then, whilst it doesn't preclude the possibility that Plato has reservations about the use of diagrams (of the sort suggested by Patterson, for instance),¹⁷ the use of diagrams is neither a defect of mathematical method, nor a failure of certain mathematicians. And likewise with respect to the use of hypotheses. The fact that mathematics does not go back to first principles and explain hypotheses is not a deficiency of mathematics. It is simply not its job to do so. This is the job of dialectic. Thus the mathematician hands over the realities he discovers to the dialectician (*Euthydemus* 290b–d).

I take this second reading to be the most compelling. A key proponent of this view is Myles Burnyeat. With respect to the use of hypotheses, Burnyeat, pointing to Plato's use of the word 'compelled' (*anankazetai/anankazomenēn*) at 510b and 511a, argues that 'the soul is compelled to start from [hypotheses] because there is no other way of doing deductive mathematics than by deriving theorems and solutions from what is laid down at the beginning'. Whilst Burnyeat's focus is on the use of hypotheses in Greek mathematics, he does also note that the use of diagrams and constructions is plausibly essential.¹⁸

(ii) The necessity of diagrams

That diagrams were, at any rate, perceived to be essential to Greek mathematics is suggested by the use of the Greek word *diagramma*. As Reviel Netz notes in *The Shaping of Deduction in Greek Mathematics*, where *diagramma* literally means 'figure marked out by lines', it is generally used in mathematical contexts to mean proposition or proof.¹⁹ This, he argues, is a reflection of the fact that 'diagrams are considered by the Greeks not as appendages to propositions, but as the core of a proposition'.²⁰ Thus Netz claims: 'Diagrams are the metonyms of propositions; in effect, the metonyms of mathematics'.²¹

Greek mathematics was geometry-heavy. On Plato's own account, plane and solid geometry form a substantial part of the *Republic* mathematical curriculum. And the use of geometry extends beyond these disciplines. Where, today, we turn geometry into arithmetic, the Greeks tackled arithmetic by means of geometry – as we see from the fact that Plato ascribes the operations of 'squaring, applying and adding and the like' to the geometers (*Rep.* 527a–b). Although Euclid is later than Plato (c. 300 BC), the general consensus is that already in Plato's time mathematics includes most of what would come to be Euclidean geometry.²²

Given the predominance of geometry, it is no surprise that the diagram should be a predominant feature of Greek mathematics, as Netz argues it is. Greek mathematical exchanges would, as a rule, he notes, be accompanied by a diagram or diagrams:²³ 'When mathematical results were presented in anything other than the most informal,

private contexts, lettered diagrams were used. These would typically have been prepared prior to the mathematical reasoning. Rulers and compasses may have been used.²⁴ As to the role of diagrams in Greek mathematics, they very clearly do a lot of work as pedagogical or heuristic aids.²⁵ But diagrams also do real work in the reasoning – whether the mathematician is using them to present or to work out a proof, whether looking at or constructing them. As Patterson notes, natural language and diagrams are part of a hybrid notational system, where both make essential contributions to proofs.²⁶ Often, the diagram supplies information that is not supplied by the collection of statements about the diagram.²⁷ (And it is worth noting that the *Republic* divided line, as a diagram, is a case in point.²⁸) In these cases, the diagram is very obviously indispensable. Arguably, too, diagrams play an indispensable role in the process of discovery. This is an argument that is increasingly made in the context of contemporary philosophy of mathematics.²⁹ In this context, of course, discussion is not confined to Euclidean geometry. But it is notable that the geometrical demonstration in Plato's *Meno* (82b–85d) is often used to illustrate the epistemologically significant role diagrams play in mathematical reasoning.³⁰ I will return to the *Meno* in the final section of this chapter.

Suffice it to say here that the use of diagrams in the *Meno*³¹ and elsewhere in Plato,³² taken together with the characterization of mathematical method in the *Republic*, suggests that Plato himself shares the perception that diagrams are a fundamental and necessary feature of mathematics – as fundamental and necessary as hypotheses. And in fact, these two features of mathematics are bound together. As Plato says at *Republic* 510b: 'In using those things which then were imitated as (themselves) images, the soul is compelled to seek from hypotheses'. The soul is compelled to seek from hypotheses in as much as it is viewing something as an image. An image is, by definition, an image of something else. So, in using something as an image, one necessarily introduces a supposition (the something else which it is *supposed* to be an image of). Using a given drawn line as an image of the diagonal is to operate with a hypothesis (that the diagonal itself exists and is like this).

3. What is the cognitive shift effected by mathematics in general?

Allowing, then, that we can identify the dianoetic method of the divided line with the mathematical method that Plato is recommending for the guardians, and that the use of diagrams is both a necessary and a predominant feature of this method, we should be predisposed to think that the use of diagrams in mathematics plays an important role in facilitating cognitive development, on Plato's account. Determining what this role might be requires us to first consider the nature of the cognitive shift effected by mathematics more generally.

(i) The intermediate cognitive state of the mathematician

Earlier in the *Republic*, at 473ff, Socrates distinguishes between sightlovers (lovers of sights and sounds) and philosophers, and their relative cognitive states. Philosophers

recognize the one, a single form; they recognize *to kalon*. They therefore have knowledge. Sightlovers, on the other hand, only recognize the many – *ta kala* – ‘beautiful sounds, colours, and shapes, and everything fashioned out of them’ (476b), and thus only have belief (*doxa*). The philosopher, because he knows what beauty is, can distinguish between beautiful and ugly things, and can distinguish between the visible and the intelligible. The sightlover cannot. Furthermore, the sightlover misidentifies the many beautiful things as beauty itself. He confuses one and many, taking the many to be the one, the likeness to be the original: ‘[He] thinks that what is like something is not like, but is the very thing to which it is like’ (476c). For this reason, the sightlover is described as being in a dream state (476c–d). The philosopher, by contrast, is ‘capable of seeing both [beauty] itself and the things which have a share of it and thinks neither that the things which have a share of it are it, nor that it is the things which have a share of it’ (476d). The philosopher does not confuse one and many, likeness and original, and is thus described as awake.

The (philosophical) mathematician is characterized as occupying the middle ground between the sightlover and philosopher. ‘I think,’ says Glaucon to Socrates, at 511c, ‘you call the cognitive state of geometers and the like *dianoia* and not *noēsis*, meaning by *dianoia* something midway between *doxa* and *nous*.’ Socrates reiterates the point at 533d–e, adding that *dianoia* has an intermediate degree of clarity.

We have already seen that Plato distinguishes *dianoia* and *noēsis* on methodological grounds. Whether or not *dianoia* and *noēsis* are also distinguishable on ontological grounds, such that the intermediate nature of *dianoia* is reflected in its objects, is a much-debated point – one fueled by Aristotle’s assertion that Plato posits intermediate mathematical objects (*Metaphysics* I.5.987b14–18).³³ I proceed here on the basis that *dianoia* does not have a distinct object.³⁴ Allowing that the ‘upper’ half of the divided line is broadly analogous with the ‘lower’ half, the implication is that *dianoia* and *noēsis* have the same object, forms, but differ in terms of the degree to which they grasp reality.

With respect to the lower half of the divided line, the objects of *eikasia* and *pistis* are essentially the same. The objects of *pistis* are sensible objects, whilst the objects of *eikasia* are images, which is to say, shadows and reflections of sensible objects. What distinguishes the two cognitive states is the clarity of their view of reality. *Pistis* apprehends reality more clearly, more directly, than *eikasia*. The person in a state of *pistis* views images of forms (sensible objects), as opposed to images of images, and recognizes that shadows and reflections of sensible objects are just that. However, he fails to recognize that those sensible objects are themselves images.

With respect to the upper half of the line, Plato describes those in a state of *dianoia* as ‘dreaming’ (*oneirōttousi*) about reality/being (*to on*) (533c). This language attributes the same sense of indirect perception or apprehension of originals to *dianoia* as is attributed to *eikasia*. The implication is that the objects of *dianoia* are reflections or shadows of forms. Whilst *eikasia* does not have a clear view of the sensible objects that cast reflections, *dianoia* does not have a clear, a direct view, of forms. Only *noēsis* does. *Dianoia* does not have a clear view of forms, because it aims at forms on the basis of hypotheses that it treats as first principles, and does not attain a truly synoptic view of forms – it does not see each form in relation to all other things.³⁵ This means *dianoia*

cannot have a true account of each form, any more than *eikasia* can have a true account of the physical object that casts the reflection.

On this reading of the divided line, what, above all, distinguishes the mathematician from the philosopher is his grasp of forms. The mathematician has some grasp of forms, such as the square itself, and the diagonal itself, but does not, as the philosopher does, fully grasp these. What, above all, distinguishes the mathematician from the sightlover is his cognitive attitude to sensible objects. Both mathematician and sightlover are described as being in a dream state, but whilst the sightlover's dream state is associated with his confusion of likeness and original, the mathematician's dream state is associated with his failure to go back to first principles (it doesn't necessarily follow from this that the mathematician *mistakes* his hypotheses for first principles, just that he treats them *as if* they were true). Notably, his dream state is not associated with his use of diagrams. The mathematician, in treating diagrams as images, treats the many as the likenesses they are.³⁶ He does not make the mistake the sightlover makes – he, like the philosopher, does not confuse the many with the one, he does not confuse likeness and original. So the mathematician has made a significant cognitive advance on the sightlover.

The intermediate cognitive state attributed to the mathematician in this way is reflected in what the mathematician does – in particular, the fact that he treats sensible objects as images – and what he *implicitly* thinks. When we are told that the mathematician is thinking about the square itself, for instance, this seems to be about his implicit, rather than explicit, view of such objects. It is not reflected in what the mathematician says:

Their language is most ludicrous, though they cannot help it, for they speak as if they were doing something and as if all their words were directed towards action. For all their talk is of squaring and applying and adding and the like, whereas in fact the whole study is engaged in for the sake of knowledge (*gnōseōs*) . . . knowledge of something always existing, and not of something coming into being at a certain time and perishing.

527a–b³⁷

(ii) Conversion and ascent

The divided line, as a static image, places the mathematician/mathematics. It indicates that the study of mathematics cultivates the intermediate cognitive state of *dianoia*. The language of conversion and ascent used of mathematics in the setting out of the mathematical curriculum, indicates that the study of mathematics is also responsible for bringing about the initial shift from a state of *pistis* to a state of *dianoia*, and for facilitating progress towards *noēsis*. It indicates, that is, that mathematics is responsible for bringing about the conversion from the sightlover's cognitive state to that of the mathematician, and for facilitating ascent towards the philosopher's cognitive state.

Conversion is a matter of being turned towards more real things (see e.g. *Rep.* 515d). It evidently requires cognitive access (up to a point) to the one, the original, that the sightlover has no access to. But, as Harte highlights, it cannot simply be a matter of

gaining cognitive access to the original, since having cognitive access to the original doesn't guarantee not confusing likeness and original. One could know Simmias, and nonetheless take a picture of Simmias to be the real Simmias. Avoiding the sightlover's confusion apparently, therefore, involves simultaneously being put in mind of the one, the original, and recognizing that what we are perceiving is a likeness of that – something which falls short of it.³⁸ This involves grasping the distinction between the visible and intelligible, likeness and original, and so developing a different cognitive attitude to sensible objects.

If mathematics is to initiate such a conversion, then clearly it must do so at the level of *pistis*, within the cave.³⁹ We have already seen that Plato posits a popular mathematics that operates at this level. Whilst he is keen to distinguish this sort of mathematics from philosophical mathematics, Plato is clear that popular mathematics can itself be educational.⁴⁰ That arithmetic at this level, simple applied arithmetic involving sensible units, is capable of prompting thought, is made explicit at *Republic* 522cff. Here, we are told that the study of arithmetic (and logistic), 'this trivial matter of distinguishing one and two and three' (522e), naturally leads to thought.

Socrates explains that certain perceptions, ones that issue in a contradictory perception, are thought-provoking. 'Number' (*arithmos*) and 'the one' belong to that class of things that impinge upon the senses with their opposite (524dff). We perceive the same thing as one, in one context, but as a many, in another context.⁴¹ To take an example from the *Parmenides*: Socrates is one, in the sense that he is one of seven men present, but also many, in the sense that he is constituted of many parts (129b–d). Since one cannot be grasped separately, or adequately, from its opposite through perception, but only confounded with many, we are puzzled as to what one is, and resort to thought to resolve the compresence of opposites that sight presents us with. 'And this (i.e. experiencing contradictory perceptions of this sort),' says Socrates, 'is how we came to use the terms "the intelligible" (*to noëton*) and "the visible" (*to horaton*)' (*Rep.* 524d).

'If this is true of the one, the same holds of all number' (525a), Socrates reasons.⁴² And since arithmetic is wholly concerned with number (525a), the study of it prompts inquiry into the one, and prompts students to distinguish it from the many. This inquiry is initiated at the level of popular arithmetic, through operations with sensible units. Subsequently, at the level of philosophical arithmetic, it is numbers understood as collections of units, where these are intelligible units that are equal to each other and admit no division into parts, that continue the conversion work.⁴³ Ultimately, the study of arithmetic leads to contemplation of the nature of numbers (525c), directing the soul upwards to discourse of the numbers themselves (*autōn tōn arithmōn*) (525d–e).⁴⁴

A striking feature of this passage is the emphasis on the physical manifestation of number/the one. Glaucon observes at 525a that it is especially the visual perception of one that provokes thought. Plato's point, I take it, is not simply that all sensible units, like army units, can be cut up such that they are at once both one and many, and so provoke thought of an intelligible one. But also, and, perhaps, especially, that the study of arithmetic can only play its part in conversion where it is, in the first instance, engaging with sensible units. Only where it engages with sensible objects can it facilitate recognition of the deficiency of sensible objects, hence recognition of the visible–intelligible distinction.

This is a point underlined in the *Phaedo*. Here, perceiving equal sticks and stones is a catalyst for the recollection process (74bff): 'It must be as a result of the senses that we obtained the notion that all sensible equals are striving to realize actual equality but falling short of it' (75a–b).

It is presumably above all for this reason that Socrates insists that the mathematical studies on the guardians' curriculum, and especially arithmetic,⁴⁵ have practical utility – that they be useful to soldiers (*Rep.* 521d). This ensures that the guardians, who will train as soldiers prior to embarking on their mathematical studies, engage in popular arithmetic, operating with sensible units (for the purposes of ordering troops, for instance), prior to engaging in philosophical arithmetic.⁴⁶

The emphasis on the order in which the mathematical subjects should be studied by the guardians might give rise to the view that arithmetic – first, at the popular level, and then, at the philosophical level – is solely responsible for conversion from *pistis* to *dianoia*. On such a view, the other disciplines might then be seen to facilitate a gradual ascent from *dianoia* to *noēsis*. And indeed the different disciplines *can* be seen to build on one another in this way. As the student of mathematics progresses through the different mathematical disciplines, he comes to perceive mathematical phenomena in contexts that gradually become more complex and broader, and thus in relation to more and more things.⁴⁷ In this way, the mathematician gradually builds towards the synoptic view of the dialectician. But Socrates indicates, too, that arithmetic is not alone in effecting conversion.⁴⁸ The suggestion is that all the subjects together effect a gradual turning around, where this is a difficult and painful process (as underlined in the cave image at 515c–516a).⁴⁹

In this case, arithmetic should rather be seen as initiating the process of conversion, instilling a low-level recognition of the visible–intelligible distinction. The other disciplines then reinforce awareness of this distinction. Geometry (plane and solid), I want to suggest, does so particularly effectively, on Plato's account, and achieves this through its use of diagrams.⁵⁰

4. How does the use of diagrams facilitate conversion and ascent?

In the famous geometrical demonstration at 82b–85d in the *Meno*, Socrates offers to show that what we call learning is actually recollection of things we already know, and proceeds to question a household servant with no knowledge of geometry on a geometrical problem: how to double a square. When, eventually, the boy discovers the theorem (that the square on the diagonal of a given square is double the area of the original square), Socrates claims that the boy has discovered this for himself. He, Socrates, has merely drawn true beliefs out of the boy that were already in him. Here, as in the *Republic*, the practice of mathematics (and for all that the boy is new to geometry, the geometry he is engaged in is nonetheless philosophical geometry) is associated with a state of dreaming: 'At this moment, those beliefs have just been stirred up in him (the boy), *like a dream*'. And, here, Plato suggests that it will be continued dialectical questioning 'about the same matters on many occasions and in many ways' (85c–d) that will lead to the awakening of the mathematician.

The process of questioning in the demonstration involves the use of diagrams. Although diagrams are not explicitly referenced, it is clear from the use of demonstratives, and the references to drawing (e.g. at 83b), that they are being used. *Phaedo* 73a–b, which refers us back to the *Meno* passage, highlights both the use of diagrams in this context, and their capacity to prompt thought:

When people are being asked questions, if someone asks the right questions, they of themselves tell everything as it is – and yet if knowledge and right account had not happened to be within them, they would not have been able to do this – for example, if someone brings them to diagrams (*diagrammata*) or something like that,⁵¹ here it most clearly affirms that this is so.

Taken together, the *Meno* and *Phaedo* suggest that the geometrical diagram is the paradigm case of a sensible object that has the power to prompt thought and recollection.

Why should Plato grant diagrams this status? Consider, for example, Euclid 1.5. The proposition here is first stated as a general claim (the *protasis* (enunciation): ‘In isosceles triangles the angles at the base are equal to one another . . .’). So it is universal. The argument, though, does not proceed by thinking about universals (isosceles triangularity). It works instead by thinking about a typical instance. The diagram is the figure you (are instructed to) draw or imagine⁵² in the setting out of the particular instance (the *ekthesis* (exposition): ‘Let ABC be an isosceles triangle having the side AB equal to the side AC . . .’). The proposition is then proved true of the particular instance. Finally, on the basis that the instance was typical, the general proposition is taken to have been proved.⁵³

The diagram in a Euclidean proof is the particular instance the geometer is thinking about, whether looking at or constructing it, and the argument is about instances (of line, triangle, even number). But the argument proceeds by logical inferences, from hypotheses (Euclidean postulates and definitions, such as the definition of ‘odd number’), and proves a universal truth (conditional on the hypotheses being true). The geometer thus uses the diagram to prove something that holds not only for the drawn/imagined triangle, the instance, but for any triangle as defined in his initial hypotheses. In so doing, he treats the diagram of an isosceles triangle, say, as representative of all isosceles triangles. As Proclus observes, geometers ‘use the objects set out in the diagram not as these particular figures, but as figures resembling others of the same sort.’⁵⁴ In the language of the *Republic*, geometers treat diagrams as images (*eikones*). They use diagrams as patterns (*paradeigmata*) to aid the study of ideal mathematical entities and relations (529d–e). Thus in the *Euthydemus* (290b–c), Plato characterizes geometers (along with arithmeticians and astronomers – presumably philosophical astronomers) – as ‘hunters’ who make use of diagrams (*diagrammata*), but are hunting mathematical truths: ‘they are not in each case diagram-makers, but discover realities (*ta onta*).’

The geometer is clear that he is using the diagram in this way, as a representation – and Plato in the *Republic* indicates as much in not ascribing the mathematician’s dream state to his use of diagrams. For one thing, a single instance is insufficient support for a universal claim. For another, the diagram is not like a picture of Simmias – it is

schematic, illustrating only the topological features of the object represented. The geometer understands that the diagram is only an instance, and only approximate – that, as Plato says in the *Seventh Letter*, ‘every one of the circles which are drawn in geometric exercises or are turned by the lathe is full what is opposite to the fifth [i.e. the form of circle], since it is in contact with the straight everywhere’ (343a). He understands that solving a problem, proving a proposition, and arriving at a geometrical truth cannot be achieved through straightforward observation or measurement of the particular diagram (and just a few attempts at trying this would confirm as much). Thus Socrates in the *Republic* observes that any geometer with experience ‘would think it absurd to examine [geometrical diagrams] seriously in the expectation of finding in them the absolute truth with regard to equals or doubles or any other ratio’ (529d–530a).

So the geometer, in using the diagram as a representation, distinguishes between the diagram, the instance, and the mathematical object as defined in his initial hypotheses, and recognizes that the former falls short – that it is approximate and suffers compresence of opposites.

Of course, using the diagram as a representation also requires the geometer to observe the similarity between diagram and mathematical object. It requires him to see the inexact diagram as the (exact) mathematical object. To put it another way, it requires him to see the inexact diagram as a picture or reminder of the (exact) mathematical object.⁵⁵ In the *Meno* geometrical demonstration, Socrates begins by asking the boy: ‘Do you know that a square figure is *like this*?’ (*toiouton estin*) (82b).⁵⁶ As Sherry puts it, Socrates’ initial line of questioning here is concerned with finding out whether the boy, as someone who is new to geometry, ‘has mastered the skill of *treating* a diagram as a square, in other words, the skill of *letting* a diagram be a square.’⁵⁷ It is a ‘skill’ honed through the practice of philosophical mathematics, and one demonstrated by the *Republic* (philosophical) mathematicians: they make use of diagrams, making their accounts about them, but think about and make their accounts for the sake of those things of which their diagrams are likenesses (510d–e). Ultimately, it is a skill that allows the geometer to access the intelligible through perception and the visible.⁵⁸

Engaging with a diagram, then, whether looking at it or constructing it, simultaneously (a) puts the geometer in mind of the ideal objects and properties instantiated there (the one, the original) (albeit that he does not fully grasp the nature of these entities); and (b) prompts the recognition that the instance(s) perceived, though similar, are distinct, and only likenesses of these, where a likeness is something that falls short of the original. In line with Harte’s reading of how the sightlover’s mistake is to be avoided, recognition of the original, and recognition of what one perceives as merely a likeness of this, occur in one and the same act.

With repeated practice, the use of diagrams will reinforce the recognition of a visible–intelligible distinction effected by arithmetic. And it will reinforce recognition of the likeness–original distinction this entails. Or, perhaps better, (rein)force. Plato repeatedly emphasizes the element of compulsion in mathematics. At *Republic* 510b and 511a (see also 511c), as we have seen, the soul is said to be compelled to start from hypotheses. At 523aff, contradictory perceptions of one and many, hence arithmetic, are said to compel thought (524c and 524e). At 526e, it is twice claimed that mathematics compels the soul to contemplate being. And the cave image emphasizes the role of

compulsion more generally in conversion and ascent (515c–e, 519c). Plato does not explicitly state that the use of diagrams compels thought, but the *Meno*, *Phaedo*, and *Republic*, together suggest that it does so in two ways. First, simply in virtue of the fact that the approximate diagram will issue in a contradictory perception: the diagram is at once circular and straight, one and many. Second, and above all, in virtue of the fact that the role of diagrams in Euclidean geometry renders them obviously representational. Diagrams therefore compel thought of the intelligible entities they represent, and compel the mathematician to recognize their similarity to and difference from these entities.

Because Euclidean geometry has grown out of a subject that is concerned with land measurement, about the physical world and sensible objects, it models the properties of the three-dimensional space that the geometer perceives around him. That Plato is very conscious of this is demonstrated in the *Meno* at 74a–76d. As Wedberg observes: ‘The concepts of Euclidean geometry are, for Plato, not variables, but certain properties and relations intimately related to our sense perception of space. In the *Meno* the general notion of geometrical figure (*schēma*) is explained in a manner which clearly shows its origin in visual perception.’⁵⁹ What is important about this (aside from the fact that it means physical diagrams can very easily be used to reason about Euclidean properties and relations) is that the geometer will associate the diagram with sensible objects – he will recognize the diagram as one among other sensible objects. Since the guardians’ study of geometry proceeds (from popular to philosophical geometry) as geometry has historically evolved,⁶⁰ this association is easily made. And it is underscored by some of the practices he engages in as a geometer: in particular, congruence proofs, where one shows that things which coincide with one another are equal to one another, by taking one figure, moving it, and placing it on another (superposition).

Recognizing that the diagram is one among other sensible objects, in this way, will prompt the further recognition that sensible objects in general are likenesses, hence that sensible objects in general are deficient. The fact that the guardians are to study astronomy (and harmonics) after geometry, is presumably, in part, because geometry – above all through its use of diagrams – will prepare them to treat stars as representations, where these are not obviously so in the way that diagrams are. As Socrates observes, the study of geometry will make for the better reception of all studies (*Rep.* 527c).

5. Conclusion

My aim here has been to demonstrate that Plato views the use of diagrams in mathematics in a highly positive light – that he sees their use as crucial and central to its capacity to facilitate conversion and progress towards knowledge of forms.

One point I have so far made very little of is the fact that the divided line image, which highlights the mathematicians’ use of diagrams, is itself a diagram – a diagram to be treated as an image (*eikon*). As such, it draws attention to key characteristics of diagrams – not least, their approximate and representational nature. Above all, though, the divided line image, as a diagram, highlights the role of the diagram in bringing out

the visible–intelligible, likeness–original distinctions. Its explicit purpose is to bring both the visible and intelligible into view at the same time (offering a synoptic view), and to highlight the differences between them, bringing the likeness–original relation to the fore.⁶¹

Notes

- 1 I presented earlier versions of this chapter at the NYU Ancient Philosophy Work in Progress Seminar, the Institute of Classical Studies Ancient Philosophy Seminar and the Oxford Ancient Philosophy Workshop. I am very grateful to audiences at these seminars for helpful comments. I am especially grateful to Hugh Benson for comments on a draft of this chapter, and to Paul Pritchard.
- 2 The exact proportions, here, are not clear.
- 3 Socrates has previously distinguished knowledge (*epistēmē*) and opinion (*doxa*). Here, knowledge encompasses intelligence and understanding, while opinion encompasses belief and imagination.
- 4 And Socrates' references to the section 'below' (*katō*), at *Rep.* 511a, and the 'highest' (*anōtatō*) section, at 511d, suggest that this is the case.
- 5 A more literal translation of 510c reads: 'In regard to these things, as if knowing [them], having made the hypotheses for themselves, they deem necessary no further account about them either for themselves or others, as being apparent to everyone'.
- 6 Plato leaves open the possibility that *dianoia* encompasses more than just mathematical reasoning. See Fine ('Knowledge & Belief in *Republic* V–VII', 2003): 107: '[Mathematics] is just one example of L3-type reasoning – Plato's moral reasoning in the *Republic* is another example of it'. The suggestion, though, is that mathematics is the most systematic way to achieve/cultivate a cognitive state of *dianoia*. See *Rep.* 518d: 'Then this turning around of the mind itself might be made a subject of professional skill (*techne*), which would effect the conversion as easily and effectively as possible'.
- 7 With respect to other sensible objects that might be used by the (philosophical) mathematician, the reference to moulding at *Rep.* 510d–e points to the use of models of solids. And astronomy uses the stars as patterns to aid them in the study of realities (529e). It also makes use of planetaria or armillary spheres (*Timaeus* 40c–d).
- 8 A notable exception is Patterson (2007). His approach, however, is importantly different from my own approach here. He considers the positive role of diagrams in the context of the *Meno* geometrical demonstration, by way of establishing what it is in the use of diagrams that Socrates/Plato is concerned about in the divided line.
- 9 See, for instance, Cross and Wooley (1964): 'First, the mathematician uses sensible images, and second, he is compelled to employ assumptions which remain unproved assumptions. These Plato regards as defects in mathematical method which the philosopher, pursuing his method of dialectic, is able to avoid' (232); and Annas (1981) asserts that mathematicians, on Plato's view, are 'complacent' and that 'mathematics has two defects compared with dialectic.' (277ff)
- 10 Klein (1968): 19–20: suggests that arithmetic is 'first and foremost the art of correct counting', while logistic is calculating, where this entails 'knowledge of the relations which connect the single numbers'. But he notes that the two terms frequently occur together and are hard to distinguish on this primary level.
- 11 See *Rep.* 528a–b and 528d.

- 12 See *Philebus* 56d: 'The mathematical arts can be divided into two kinds: that of the many (*tōn pollōn*), and that of the philosophers'. The distinction is essentially one between applied and theoretical mathematics, where the latter is more accurate.
- 13 As to how to understand this, Burnyeat (2000): 15, n.18, suggests that 'problem' here should not be translated 'in such a way as to *confine* Platonic astronomy and harmonics to problems in the technical sense' – i.e. constructions, as distinct from theorems. See also Benson (2012): 188.
- 14 For instance, units *as such*, as opposed to units of oxen or army.
- 15 See esp. Benson (2012); also Benson (2010).
- 16 Benson (2012) allows that 'it is difficult to see why the use of sensible diagrams and constructions should be thought to be problematic' (194), but suggests that inappropriate use of ordinary sensible objects might roughly consist in 'mistaking contingent or artificial consequences of one's hypothesis for genuine or natural consequences . . . For example, one might take the hypothesis that the circumference of a circle is equivalent to twice the product of π and the radius . . . to be refuted by measuring the circumference of a given circle' (194–195). He concedes that no accomplished mathematician would do this, but argues that unaccomplished ones, beginners of geometry, might be inclined to make this mistake – these geometers must learn not to allow accidental features of the diagram too much weight, and to recognize the essential features of the diagram. Benson's view is compatible with the view that diagrams are a fundamental feature of Greek mathematics, and that they play a key role in conversion, as I argue here, but he takes the line that it is because they are fundamental that Plato emphasizes their misuse in *dianoia*.
- 17 Patterson (2007) suggests that Plato is worried 'the use of diagrams reflects, promotes, and to a large extent constitutes a thinking of the abstract in the sensible and particular . . . The use of diagrams, if not understood from the larger perspective of Platonic metaphysics and theory of knowledge, condemns the practitioner to ignorance of the true foundations both of mathematical truth and mathematical cognition' (2). But he is also clear that it does not follow from this that Plato would want to ban the use of diagrams or reform the mathematical method (30, 33).
- 18 Burnyeat (1987): 218–219; and (2000): 37–38.
- 19 Netz (1999): esp. 35–40.
- 20 Netz (1999): 35.
21. Netz (1999): 40.
22. See Heath (1921): esp. 216–217. Heath concludes that 'there is probably little in the whole compass of the *Elements* of Euclid, except the new theory of proportion due to Eudoxus and its consequences, which was not in substance included in the recognized content of geometry and arithmetic by Plato's time' (217). See also Burnyeat (2000): 24: 'Euclid's *Elements* incorporates much previous work, from two main sources: first, earlier *Elements* by Leon and Theudius, both fourth-century mathematicians who spent time in the Academy, and, second, the works of Theaetetus and Eudoxus, two outstanding mathematicians with whom Plato had significant contact. If we could read the mathematics available at the time Plato wrote the Republic, a good deal of it would look like an early draft of Euclid's *Elements*'.
- 23 Netz (1999): 14.
- 24 Netz (1999): 19. Netz suggests that diagrams would have been prepared on a range of media, including dusted surfaces, wax tablets, papyri and whiteboards (15–16). He notes that his own efforts to draw diagrams in the sand were an unmitigated disaster – which raises questions regarding the standard interpretation of the geometrical

demonstration in the *Meno*, on which Socrates draws diagrams in the sand as he speaks.

- 25 See Patterson (2007) for examples.
- 26 Patterson (2007): 14.
- 27 See Patterson (2007): 14–17. And Netz (1998) argues that: 'Greek mathematics relies upon diagrams in an essential, logical way. Without diagrams, objects lose their truth-value. *Ergo*, part of the content is supplied by the diagram, and not solely by the text. The diagram is not just a pedagogic aid, it is a necessary, logical component' (34).
- 28 It emerges through construction of the diagram that the two middle sections of the line are equal in size. This is not something that is specified in the text.
- 29 See, for instance, Brown (2008): esp. chs. 3 and 12; and Giaquinto (1993) and (2008).
- 30 See esp. Giaquinto (1993).
- 31 Both in the geometrical demonstration, and also at 86dff.
- 32 For instance, at *Theaetetus* 147d–148b, and esp. at *Timaeus* 40c–d: 'To describe all this (the movements of the heavens) without an inspection of models (*mimēmatōn*) of these movements would be labour in vain' (40d).
- 33 For the view that *dianoia* and *noēsis* are *not* distinguishable on ontological grounds, see, for instance, Pritchard (1995): 108–111. See also Fine ('Knowledge & Belief in *Republic* V–VII', 2003): 99–104 who argues against an objects analysis and for a contents analysis of the *Republic* divided line. For the view that *dianoia* and *noēsis* are distinguishable on ontological grounds, and that the objects of mathematics are intermediates, see, for instance Annas (1988): esp. 251, Denyer (2006): 303, and Burnyeat (2000): 33ff.
- 34 It is clear that Plato *should* posit intermediate mathematical entities – it is not possible to do mathematics with forms – and there are indications outside of the *Republic* divided line passage that he does (see esp. *Philebus* 55c–59c). But I see no compelling evidence for this view within the divided line passage. *Rep.* 534a suggests that Plato is deliberately side-stepping discussion of mathematical objects, in order to bring out the cognitive differences between mathematician and philosopher, by means of the image–original relation.
- 35 Cross and Woosley (1964): 238 argue that the objects of *dianoia* and *noēsis* differ in as much as the forms that *noēsis* grasps, unlike the forms that *dianoia* grasps, are grasped in all their interconnectedness and in the light of the Good.
- 36 Some mathematicians do make the sightlover's mistake – for instance, the astronomers criticized for focusing their attention on the stars, and not what is beyond. But these are the mathematicians of the day, not the philosophical mathematician.
- 37 Or: '... knowledge of something which is always the case, and not of something which is sometimes the case and sometimes not'.
- 38 See Harte (2006). See esp. 37: 'On my reading, if one recognizes the equality of perceptible equals in the manner that my interpretation of the crucial experience requires, recognition of the form, and recognition of what one perceives as merely a likeness of the form must occur in one and the same act'.
- 39 See Pritchard (1995): 106; also Gregory (1996): 453.
- 40 Plato distinguishes the use of arithmetic and logistic in war from its use for the purposes of buying and selling (*Rep.* 525b–c), thus distinguishing between educational practical pursuits and purely practical ones. Rosen (2005): 290–291 suggests that this distinction is made on the basis that moneymaking is base. At 536d–537a, the practice of mathematics is advocated for young children. See also *Laws* 819.

- 41 There are many different interpretations of this famous passage, but on one possible reading Plato is identifying relative properties, such as large and small, and one, also, as things that impinge upon the senses together with their opposite because they are 'incomplete' predicates, context-dependent.
- 42 The point seems to be that all numbers attached to visible and tangible bodies suffer the compresence of one and many, hence provoke thought. Presumably because every number is made up of ones. So three, say, is one number ('one' in a series), but also three units, and one and many in this sense (see Craig 1994: 282). Thus where we perceive three oxen, we see *a* three, but also three units, hence one compresent with many. And so it is with all numbers, that they appear no more one than its opposite, many.
- 43 See e.g. Pritchard (1995): 65: 'a Greek number must be some definite plurality (collection) of units (monads) where the unit is either some physical object or an abstract unit'.
- 44 Some commentators (see, for instance, Cross and Woosley 1964: 237) have taken this to mean that the study of arithmetic prompts contemplation of *forms of numbers* – *all* numbers, not just One – so the form of Two and the form of Three etc. If, however, all numbers suffer the same compresence of one and many, then even if there *are* forms of all numbers (*Phaedo* 101b–c, for instance, gives the impression that there are), observation of this compresence of one and many will always provoke contemplation and discourse of the form of One, at least in the first instance. In this case, Socrates must mean that we are prompted to contemplation and discourse of 'the nature of numbers', where it is in the nature of every number to be one.
- 45 See esp. *Rep.* 522c–e; also 525b and 526c.
- 46 The requirement for practical utility is often dismissed as insignificant, but is surely crucial for the purposes of conversion. Shorey (1935) 148, n.(c), for instance, notes: 'This further prerequisite of the higher education follows naturally from the plan of the *Republic*; but it does not interest Plato much and is, after one or two repetitions, dropped'. See also Burnyeat (2000): 10–11: 'Plato is not serious about justifying the study [of arithmetic] on grounds of its practical utility'.
- 47 Plane geometry is concerned with two dimensions, solid geometry, three dimensions, and astronomy, solids in revolution (*Rep.* 528a–b, 528d–e). Harmonics, like astronomy, is a study of motion – 'as the eyes fasten on astronomical motions, so the ears fasten on harmonic ones' (530d). Understanding something in three dimensions requires recognizing how it stands in relation to more other things, for instance depth, than if it were two-dimensional. And understanding a three-dimensional object in motion will require recognizing how it stands in relation to even more things, such as time.
- 48 See e.g. *Rep.* 526e, 527b, 529a, 533d. Socrates' insistence on the practical utility of mathematical subjects other than arithmetic, might also be taken to demonstrate this – plane geometry applies to the conduct of war (526c); and astronomy is 'serviceable not only to agriculture and navigation, but still more to the military art' (527d).
- 49 See Craig (1994): 285: 'The sequence of five studies . . . is not simply and exclusively an anabasis ("ascent"), although rhetorically that is the dominant impression created. Partly this is a carry-over from the Cave allegory . . . This impression is reinforced by the natural pedagogic progression of the studies themselves . . . But the progression of study from arithmetic to harmony doesn't bring one ever closer to Being, nor is it said to'.
- 50 One might also make the case that popular geometry – geometry applied to the sensible world, e.g. for building houses or organizing troops – already plays a

conversion role, just as popular arithmetic does. But I want to argue here that it is especially the use of diagrams that effects conversion.

- 51 For instance, models of solids, where these have the same function as diagrams.
- 52 We don't have to cash out the use of visible objects in terms of physical diagrams: we could just as well be *imagining* a particular triangle – this will involve imagining a triangle to be a particular size, say.
- 53 See Mueller (1981): 13: 'The *protasis* is formulated without letters to make the generality of what is being proved apparent. The *ekthesis* starts the proof, but, before the proof is continued, the *diorismos* insists that it is only necessary to establish something particular to establish the *protasis*. When the particular thing has been established, the *sumperasma* repeats what was insisted upon in the *diorismos*'.
- 54 *A Commentary on the First Book of Euclid's Elements*, trans. G. R. Morrow (1970): 207.
- 55 Giaquinto (1993) 90: 'The diagram must look similar to something which appears exactly square. Seeing the diagram as a square involves both seeing it and observing this similarity of appearance'.
- 56 Although the *Meno* demonstration is not set out like a Euclidean proposition, as Patterson notes (2007) 5ff, it roughly speaking constitutes a Euclidean production proof.
- 57 Sherry (2008): 64.
- 58 See Giaquinto (1993) on the *Meno* demonstration: 'Although the process includes bits of deductive sentential reasoning, the use of diagrams in this process is clearly not a superfluous adjunct to a proof (a valid sequence of sentences), since no proof of the theorem was followed or constructed. On the other hand, the use of diagrams was not empirical: the visual experience that resulted from the use of diagrams was not used as a source of observational evidence for this or that proposition. In this case vision was a means of getting information about things that were not before one's eyes. Seeing the diagram as a geometrical figure of a certain sort, seeing parts of it as related in certain geometrical ways and visualizing motions of the parts, enabled us to tap our geometrical concepts in a way which feels clear and immediate.' See also Brown (2008), who argues that diagrams (in the context of contemporary mathematics) can be 'windows to Plato's heaven' (40).
- 59 Wedberg (1955): 47.
- 60 I owe this point to Andy Gregory.
- 61 Brumbaugh (1968) observes that Plato in general makes prominent use of diagrams to represent the intelligible-visible or knowledge-opinion relationship, suggesting that he sees diagrams as naturally fitted to bring out this relationship.

References

- Annas, J. (1981). *An Introduction to Plato's Republic*. Oxford: Oxford University Press.
- Benson, H. H. (2010). 'Plato's philosophical method in the *Republic*: The divided line (510b-511d)'. In M. McPherran (ed.), *Cambridge Critical Guide to the Republic*. Cambridge: Cambridge University Press, pp. 188–208.
- Benson, H. H. (2012). 'The problem is not mathematics, but mathematicians: Plato and the mathematicians again'. *Philosophia Mathematica* (3)20: 170–199.
- Brown, J. R. (2008). *Philosophy of Mathematics*. London and New York: Routledge.
- Brumbaugh, R. S. (1968). *Plato's Mathematical Imagination: The Mathematical Passages in the Dialogues and Their Interpretation*. Bloomington, IN: Indiana University Press.

- Burnyeat, M. F. (2000). 'Plato on why mathematics is good for the soul'. *Proceedings of the British Academy* 103: 1–81.
- Burnyeat, M. F. (1987). 'Platonism and mathematics: A prelude to discussion'. In A. Graeser (ed.), *Mathematics and Metaphysics in Aristotle*. Bern: Haupt, pp. 213–240.
- Craig, L. H. (1994). *The War Lover: A Study of Plato's Republic*. Toronto: University of Toronto Press.
- Cross, R. C. and Woosley, A. D. (1964). *Plato's Republic: A Philosophical Commentary*. New York: Macmillan.
- Denyer, N. (2007). 'Sun and line: The role of the good'. In G. R. F Ferrari (ed.), *The Cambridge Companion to Plato's Republic*. Cambridge: Cambridge University Press, pp. 284–309.
- Fine, G. (2003). *Plato on Knowledge and Forms*. Oxford: Oxford University Press.
- Harte, V. (2006). 'Beware of imitations: Image recognition in Plato'. In F. G. Hermann (ed.), *New Essays on Plato*. Swansea: Classical Press of Wales, pp. 21–42.
- Heath, T. (1921). *A History of Greek Mathematics*, vol. 2. Oxford: Clarendon Press.
- Giaquinto, M. (1993). 'Diagrams: Socrates and Meno's slave'. *International Journal of Philosophical Studies* 1: 81–97.
- Giaquinto, M. (2008). 'Visualizing in mathematics'. In P. Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 22–64.
- Gregory, A. (1996). 'Astronomy and observation in Plato's Republic'. *Studies in History and Philosophy of Science* 27(4): 451–471.
- Klein, J. (1968). *Greek Mathematical Thought and the Origin of Algebra*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Mueller, I. (1981). *Philosophy of Mathematics and Deductive Structure in Euclid's Elements*. New York: Dover Publications.
- Netz, R. (1998). 'Greek mathematical diagrams: Their use and their meaning'. *For the Learning of Mathematics* 18(3): 33–39.
- Netz, R. (1999). *The Shaping of Deduction in Greek Mathematics: A Study in Cognitive History*. Cambridge: Cambridge University Press.
- Patterson, R. (2007). 'Diagrams, dialectic, and mathematical foundations in Plato'. *Apeiron* 40: 1–33.
- Pritchard, P. (1995). *Plato's Philosophy of Mathematics*. Berlin: Academia-Verlag.
- Proclus (1970). *A Commentary on the First Book of Euclid's Elements*, trans. G. R. Morrow. Princeton, NJ: Princeton University Press.
- Rosen, S. (2005). *Plato's Republic*. New Haven, CT: Yale University Press.
- Sherry, D. (2008). 'The role of diagrams in mathematical arguments'. *Foundations of Science* 14(1–2): 59–74.
- Shorey, P. (1935). *Plato: The Republic, Books VI-X*. Cambridge, MA: Harvard University Press.
- Wedberg, A. (1955). *Plato's Philosophy of Mathematics*. Stockholm: Almqvist & Wiksell.

Generality

Nils Kürbis

1. Introduction

In Plato's *Sophist* the Stranger from Elea gives Theaetetus good advice: 'In fact, my friend, it's inept to separate everything from everything else. It's the sign of a completely unmusical and unphilosophical person' (259d–e). It is advice I can imagine Keith Hossack – my *Doktorvater*, a completely musical and philosophical person – giving his student. It was a privilege to study with a philosopher whose thought exemplifies rigour, intellectual honesty and an encompassing philosophical vision, and I am grateful for our ongoing discussions and friendship. The question 'Would Keith let me get away with this argument?' is still a driving force in getting clear in my philosophical thinking. This chapter concerns an aspect at the heart of Hossack's philosophy: realism, or a theory of why it's inept to separate everything from everything else.¹

Hossack's *The Metaphysics of Knowledge* contains a sustained argument that there are *negative facts*, facts that contain the universal *negation* in principal position (Hossack 2007: Sections 2.2–2.6). Negative facts are required to distinguish sense from nonsense, a distinction in turn required in the definition of particulars, universals and the arity of universals. Negation is a logical constant. The universal *negation* is the referent of the symbol for negation in logic. Complex facts are facts that contain referents of logical symbols. Consequently, there are complex facts. But are there any other complex facts besides those containing *negation*? The sections arguing for the existence of negative facts are followed by a much briefer section considering whether there are complex facts containing universals referred to by two other expressions of logic, conjunction and universal quantification. Hossack accepts that besides the negative facts there are also complex facts containing the universal *conjunction* and those containing the universal *generality*. These are the conjunctive and the general facts (Hossack 2007: Section 2.7). But Hossack gives no argument comparable to the intricate line of reasoning that concludes that there is a universal *negation*. Hossack arrives at his conclusion after having found the theory of deduction of Wittgenstein's *Tractatus* wanting. Its positive support is rather brief. I will here only consider generality. Hossack bases his acceptance of general facts on 'broader metaphysical considerations', of which 'by far the most important is that the existence of general facts is required if

we are to give a realist account of laws of nature: for a law is a general fact' (Hossack 2007: 72). While the existence of negative facts is concluded from an argument at the very foundations of Hossack's metaphysics, the existence of general facts is presented as a requirement of more remote developments of his views.

In this chapter, I will offer Hossack three arguments for his conclusion, which are of a foundational character. They draw heavily upon the views of Russell, one of Hossack's philosophical heroes. Two of the arguments appeal to the nature of universals and one to the nature of logical inference.

For most of this chapter I will restrict consideration to generality of first-order, the generality that corresponds to quantification in first-order predicate logic, to avoid complications with paradoxes such as Russell's and the Liar. Russell did not accept that it is possible to quantify over absolutely everything, and the issues I am going to address already arise in the case of quantification over particulars. At one point there will be occasion to mention generality of higher-order. Hossack rejects any such distinction of orders and avoids the threat of paradox by other means. An exposition of the details would go too far here, but I see no reason why the arguments given here may not be adapted to his framework.

Russell's views on generality are intriguing and merit a historical investigation, which will be the subject of another paper (Kürbis (ms.)). They constitute one of Russell's famous changes of mind. The view Hossack attributes to Wittgenstein, according to which there are no general facts and general statements are made true by their instances, was also held by Russell in the first edition of *Principia Mathematica* (Russell and Whitehead 1910: 47f). In *The Philosophy of Logical Atomism*, Russell had moved to accepting that there are general facts. Whereas in the first edition of *Principia*, there are evidently no general facts, only general judgements, in *The Philosophy of Logical Atomism*, Russell takes it as evident that there are and even goes as far as saying that it cannot be doubted that there are general facts (Russell 1919b, Lecture V, 200). Russell does not, on the other hand, say much about the nature of general facts. This is a difference between Russell and Hossack we should note right from the start. In *The Philosophy of Logical Atomism*, Russell doubts that there are molecular facts, facts of the kind Hossack calls 'complex' (Russell 1919a, Lecture III, 39, 42), but accepted that there are negative facts. Around the same time, Russell also explicitly rejected the view that negation refers to something, so negative facts are not molecular. The distinction between negative and positive facts is one of form or quality, not of constituents. 'It must not be supposed that the negative fact contains a constituent corresponding to the word "not." It contains no more constituents than a positive fact of the correlative positive form. The difference between the two forms is ultimate and irreducible' (Russell 1919c: 4). Russell is less explicit about what distinguishes general facts. Hossack, by contrast, is: they contain the universal *generality*. In the light of the argument to be given in the next section, it is somewhat surprising that Russell did not draw Hossack's conclusion: it is suggested by Russell's view on the nature of universals, and at one point Russell even mentions 'abstract, logical universals' of which we may have knowledge by acquaintance (Russell 1912: 171). Be that as it may, Russell's notion of general fact is rather different from Hossack's: Hossack's general facts are molecular or complex facts, Russell's apparently are not. It may well be that Russell had in mind

that general facts, and possibly others one might be inclined to call molecular or complex, are really further forms of facts. But to argue this point is a matter for another occasion.²

2. Resemblance between universals. I

Russell's characterization of the nature of universals and how they are distinguished from particulars allows for a swift argument for the existence of general facts and Hossack's conclusion as to their nature. The distinction between universals and particulars is explained in an especially clear and striking fashion in *The Problems of Philosophy*. Russell reflects on the characteristics of words such as 'whiteness' and 'justice':

The word will be applicable to a number of particular things because they all participate in a common nature or essence. [...] We speak of whatever is given in sensation, or is of the same nature as things given in sensation, as a particular; by opposition to this, a universal will be anything which may be shared by many particulars, and has those characteristics which, as we saw, distinguish justice and whiteness from just acts and white things.

Russell 1912: 143, 145

Some things resemble each other, others don't, and this is a matter of fact, not of opinion or perception. As Hossack puts it: 'Things resemble because they literally have something in common, namely a universal. [...] Resemblance is an objective matter: it is not just that some things strike us humans as similar by human standards; some things really are similar, whether they strike us as similar or not' (Hossack 2007: 35). A universal can be shared by different things, the things that resemble each other. Russell also points out that at least *resemblance* must be a universal, but that an attempt to do away with all other universals and keep only *resemblance* is doomed to failure (Russell 1911: 8ff). Things do not merely resemble each other, but do so in certain respects, and these respects are the universals. A universal is an aspect of resemblance.³

Another point that may be added is that objects not only resemble each other in certain respects, but may also fail to do so. In such a case, they do not share these universals. Assuming Leibniz's Law, objects that are different do not resemble each other in certain respects.

Russell focusses on characteristics that may be shared by particulars. Socrates, Plato and Aristotle resemble each other in being men. Bucephalus and Incitatus resemble each other in being horses. Once we have accepted that there are universals, we may note that there are relations between them. General facts are of this kind. Russell gives as an example the fact that all men are mortal (Russell 1919c, Lecture V, 200f), which consists in a relation between the universals *man* and *mortality*. Another example is the fact that all horses are hoofed, which consists in a relation between the universals *horse* and *hoofed*. We may now note that these pairs of universals resemble each other. The pair *man* and *mortality* has something in common with the pair *horse* and *hoofed*: they

resemble each other in that the first is subsumed under the second. Given Russell's explanation of the nature of universals, we should therefore conclude that there is a universal corresponding to what in *Principia Mathematica* is called *formal implication* (Russell and Whitehead 1910: 21f). The same relation holds between the two pairs.

Russell's discussion of $2 + 2 = 4$ in *The Problems of Philosophy* mentions further characteristics that may be shared by universals:

In the special case of 'two and two are four,' even when we interpret it as meaning 'any collection formed of two twos is a collection of four,' it is plain that we can *understand* the proposition, *i.e.* we can see what it is that it asserts, as soon as we know what is meant by 'collection' and 'two' and 'four.' It is quite unnecessary to know all the couples in the world: if it were necessary, obviously we could never understand the proposition, since the couples are infinitely numerous and therefore cannot all be known to us. Thus although our general statement *implies* statements about particular couples, *as soon as we know that there are such particular couples*, yet it does not itself assert or imply that there are such particular couples, and thus fails to make any statement whatever about actual particular couples. The statement made is about 'couple,' the universal, and not about this or that couple.

Thus the statement 'two and two are four' deals exclusively with universals, and therefore may be known by anybody who is acquainted with the universals concerned and can perceive the relation between them which the statement asserts.

Russell 1912: 163f

Couple is a universal, but it is not a characteristic of particulars like Socrates and Plato. It might be a characteristic of pluralities or collections, but if we try to attribute it to any of the entities Russell explicitly accepted at this point, and he was of course sceptical whether classes 'really' exist, then there is little choice but to attribute it to universals: two things are a couple if they share a universal shared only by them. *Having exactly two instances* is a universal of universals, a characteristic shared for instance by the universals *planet closer to the sun than Earth* and *planet between Venus and Jupiter*.

Thus having accepted that there are aspects of resemblance, *i.e.* universals, we can further observe that there are relations between them and that there are aspects of resemblance between the universals themselves. The universals *mortality*, *rationality* and *animality* resemble each other in the respect of *being had by every human*. *Having exactly two instances* is an aspect of resemblance between the universals *planet closer to the sun than Earth* and *planet between Venus and Jupiter*. Generality is also something in respect of which universals can resemble each other: *being universally instantiated* is a property shared by every universal had by every particular, such as *being a material object that is extended*, *being mortal if human*, *being hooved if a horse*. We should thus side with Hossack and conclude that there is a universal *generality*.

We may call such resemblances higher-order resemblances, but resemblances they are nonetheless. Given Russell's explanation of what a universal is, there is no reason not to accept that these higher-order relations or properties are universals. The

generality of the examples given is first-order, but a statement of the resemblances between them involves expressions of higher-order logic. *Having exactly two instances* means that there are x, y such that Fx, Fy and x and y are different and anything that is F is either x or y . That two universals F, G resemble each other in this respect is expressed in Fregean fashion by the sentence stating that there is a one-to-one correspondence between the F s and the G s, i.e. by the corresponding instance of Hume's Principle, which requires second-order quantification.

According to Hossack and Russell, universals combine with particulars and other universals to form facts (Hossack 2007: 45ff). The present line of argument concludes that there is a universal *generality*. This universal is a constituent of facts. It combines with such properties as *being mortal if human*, *being hoofed if a horse* and *being a material object that is extended* to form facts corresponding to 'All humans are mortal', 'All horses are hoofed', 'Every material object is extended'. These are what Hossack calls 'general facts'.

Another characteristic that is often attributed to universals is that, contrary to particulars, they can be in many places at the same time. If it makes sense to say this, it also makes sense to say that a universal occurs only once. Adopting terminology from logical syntax, if the universal *generality* is the principal universal in a fact, then the fact is a general fact; if *negation* is the principal universal, it is a negative fact. A fact that corresponds to a double negation would then count as a negative fact, which may not satisfy everyone, but as this is not the topic of the present chapter, I shall not attempt to sort out this issue.

This line of argument to the conclusion that there is a universal *generality* is not explicitly taken by either Russell or Hossack. Given Russell's discussion of couples, this is somewhat surprising. It may be due to a reluctance to accept that logical constants refer. But the line of argument follows their shared outlook on the nature of universals closely. As will be discussed later, some such line has in fact been taken by some metaphysicians, although I will argue that further aspects of their views mean that they are less satisfactory than those of Russell and Hossack.

3. Resemblance between universals. II

According to Russell, existence is a property, albeit not a property of individuals: 'Existence is essentially a property of a propositional function. It means that the propositional function is true in at least one instance. [...] Existence is a predicate of a propositional function, or derivatively of a class' (Russell 1919b: Lecture V, 195f). This is also fundamentally Frege's view: existence is a concept of concepts (Frege 1884: §53). Not wishing to go into either Russell's views on propositional functions or Frege's on concepts, we can, instead, build on material from the last section to form an alternative view on what existence is.

Socrates, Plato and Aristotle are human. Bucephalus and Incitatus are horses. The universals *human* and *horse* have something in common: they are instantiated. There is something that has that property or something with that property exists. Russell explicitly accepted that there are existential facts (Russell 1919b: Lecture V, 200). But he

shied away from concluding that there is a universal *existence*. Given his explanation of the nature of universals, there may be little reasons for this reluctance. Instantiated universals resemble each other in this respect. Thus there is a universal that is this resemblance. We may call it 'existence'. It is a higher-order universal, shared by some universals. Existential facts are those in which the universal *existence* is combined with other universals, such as *human* or *horse*.

Some properties are distinct from *human* and *horse* in this respect: *unicorn* and *winged horse* are not instantiated. But this means that everything fails to be a unicorn and is other than a winged horse. Russell accepted that there are negative facts without committing to the view that the symbol for negation refers to a constituent of facts. Arguably, however, objects that lack a property resemble each other in the respect of not instantiating this property. Such facts have something in common. This provides further support for Hossack's conclusion that there is the universal *negation*. Either way, existential and negative facts together yield universal facts. Conversely, if there are negative facts and general facts, there are existential facts.

Some philosophers may now feel that surely the account proves too much. There are too many universals, if all that is required for the existence of a universal is some aspect of resemblance. Some philosophers may, for instance, prefer to say that existential propositions are made true by their instances. If Socrates is human, then there are humans. The first fact suffices to ensure the truth of the sentence expressing the latter, and there is no need for existential facts in addition.

Two responses can be made here. Positions according to which properties are 'abundant' have been argued for.⁴ Such a view, however, does not cohere well with Russell's and Hossack's characterization of the nature of universals. A theory according to which universals are aspects of resemblance will probably not accept that they are 'abundant': not everything resembles everything else in the substantial fashion countenanced by the theory. Such a theory will rather take to heart Socrates' remarks in the *Phaedrus*. There are, says Socrates,

two kinds of things the nature of which it would be quite wonderful to grasp by means of a systematic art. [...] The first comes in seeing together the things that are scattered about everywhere and collecting them into one kind [...]. The other] is to be able to cut up each kind according to its species along its natural joints, and to try not to splinter any part, as a bad butcher might do.

Phaedrus 265d–e

The first aspect is to collect things together according to their resemblances, the second to look more deeply into what explains the resemblances according to the natures of things. Not every resemblance needs to point to the existence of an unanalysable universal. It suffices that any resemblance can be explained in terms of or reduced to some universals. This line of thought would not exclude the possibility that nature allows for some abundance in her offerings.

Russell happily accepted that there are universal, existential and negative facts, even though one of the former two would appear to be redundant in the presence of the other and negative facts.⁵ It would be natural to call an existential fact also a general

fact, as does Russell, albeit one of a different kind than a universal fact. Hossack, by contrast, accepts only that there is the universal *generality*, not also the universal *existence*, and only calls facts containing the former in principal position 'general facts'. Existential facts are analysed in terms of generality and negation.⁶

4. The logic of generality

While the arguments of the previous two sections were not put forward by Russell, there is an explicit argument for the existence of general facts in *The Philosophy of Logical Atomism*. Here Russell turns away from the treatment of generality of the first edition of *Principia Mathematica*, the same that Hossack finds wanting. General judgments are said to collect together elementary judgments and they are made true by their instances:

The judgment ['all men are mortal'] does not correspond to one complex, but to many, namely 'Socrates is mortal,' 'Plato is mortal,' 'Aristotle is mortal,' etc. [...] Our judgment that all men are mortal collects together a number of elementary judgments [namely that Socrates is mortal, that Plato is mortal, that Aristotle is mortal, etc.]. It is not, however, composed of these, since (*e.g.*) the fact that Socrates is mortal is no part of what we assert, as may be seen by considering the fact that our assertion can be understood by a person who has never heard of Socrates. In order to understand the judgment 'all men are mortal,' it is not necessary to know what men there are.

Russell and Whitehead 1910: 47

Russell's argument for the existence of general facts is based on the observation that no universal proposition follows from exclusively particular premises: any such inference is invalid. From elementary propositions about individual men's mortality, 'Socrates is a man that is mortal,' 'Plato is a man that is mortal,' 'Aristotle is man that is mortal' and so on, one cannot infer 'All men are mortal' unless the premise 'All men are amongst those I have enumerated' is also given, which is a general proposition. 'You never can arrive at a general proposition by inference from particular propositions alone. You will always have to have at least one general proposition in your premise' (Russell 1919c, Lecture III, 198f). Certain logical inferences require general propositions for their validity. Thus they cannot be reduced to their instances. General propositions must therefore state general facts, rather than collecting together elementary facts. Accordingly, Russell concludes that there are irreducibly general facts:

I do not think one can doubt that there are general facts. It is perfectly clear, I think, that when you have enumerated all the atomic facts in the world, it is a further fact about the world that those are all the atomic facts there are about the world, and that is just as much an objective fact about the world as any of them are. It is clear, I think, that you must admit general facts as distinct from and over and above particular facts. The same thing applies to 'All men are mortal.' When you have

taken all the particular men that there are, and found each one of them severally to be mortal, it is definitely a new fact that all men are mortal; how new a fact, appears from what I said a moment ago, that it could not be inferred from the mortality of the several men that there are in the world.

Russell 1919b, Lecture V, 200

A mere list of mortal men separates each item from every other, as the Stranger from Elea might say, and falls short of the truth about all men: that they are all mortal.

This may still allow for some leeway in what general facts there are. It may, for instance, suffice to list all atomic truths and add the clause that these are all of them. This is the position taken in the introduction to the second edition of *Principia Mathematica*:

Given all true atomic propositions, together with the fact that they are all, every other true proposition can theoretically be deduced by logical methods. That is to say, the apparatus of crude fact required in proofs can all be condensed into the true atomic propositions together with the fact that every true atomic proposition is one of the following: (here the list should follow).

Russell and Whitehead 1997: xv

It may also suffice to list all atomic propositions and then to add that the particulars named in them are all there are. Formally, this is the more straightforward option, at least in the finite domain, as it allows us to remain firmly first-order. Suppose there are only two objects in the universe, a and b name them, and each of them has the property F . Then the inference from premises (1) Fa and (2) Fb to the conclusion $(x)Fx$ is invalid. If we add a further premise to the effect that a and b are the only objects there are, i.e. (3) $(x)(x = a \vee x = b)$, we can argue as follows: by universal quantifier elimination, from (3) $x = a \vee x = b$; if the first, then by (1) Fx , if the second, then by (2) Fx , hence by disjunction elimination Fx ; Fx was derived from premises not containing x free, hence by universal quantifier introduction, $(x)Fx$. This pattern of argument can be extended to any finite domain of objects.⁷

It may be that all cases of generality can be reduced to this one kind: once we have listed all humans and observe that each one of them is mortal, or all the facts about human mortality, and add that these are all of them, we may infer that all humans are mortal. But whether there is only one kind of general fact or several, there are then irreducibly general facts.

On the other hand, if the truth of elementary propositions is defined in terms of the facts to which they correspond (Russell and Whitehead 1910: 43f), and facts are explained as containing universals (Russell 1918, Lecture II, 513), a less austere route is more natural. Although it is not necessary to know who Socrates was in order to understand that all humans are mortal, it is necessary to know the universals *human* and *mortality*, as pointed out in *The Problems of Philosophy* (Russell 1912: 166). There it is also pointed out that certain truths consist purely in the relations between universals. Those that give rise to *a priori* knowledge are amongst them (Russell 1912: 162). Our evidence for the truth of 'All humans are mortal' comes from our observation

of individual instances of human mortality, but its meaning is one that deals only with universals. Russell was never in doubt that there are facts concerning the relations between universals. In particular, there are those about which universals subsume which others. Facts concerning only the relations between universals contain no particulars and are therefore general. Thus it is plausible that there are rather more general facts than just those concerning enumerations of the facts or objects that there are.

Russell observes that if there are general facts concerning the subsumption of one universal under another, it may be hard to resist the conclusion that there are hypothetical facts that link particular instantiations of those universals, such as the fact that Socrates is mortal if a man (Russell 1919b, Lecture V, 201).⁸ If there were, he would probably have to admit that there is a universal *hypotheticality* or a further form of facts.

Russell resisted the conclusion that the subsumption of one universal under another is also a universal. This may again have something to do with worries concerning logical constants. Russell does not say anything more about the nature of general facts. The first argument given here, however, settles the issue.

5. Comparison with other metaphysicians' views on generality

Let's compare Russell's and Hossack's views with those of two other contemporary defenders of general facts, Armstrong and Hochberg.

According to Armstrong, if there is a number of facts and these are all of them, then they stand in the '*alling* or *totalling* relation' (Armstrong 1997: 199). Armstrong's argument for the existence of such a relation is reminiscent of that of Section 2 above: 'It can hardly be denied that such a relation exists [. . .]. It looks to be exactly the same relation in each case, to be a true universal if you accept universals' (Armstrong 2004: 73). Totality facts present limits to what there is:

If it is true that a certain conjunction of states of affairs is all the states of affairs, then this is only true because there are no more of them. If there are more, then the proposition is not true. That there are no more of them must then be brought into the truthmaker. But to say that there are no more of them is to say that they are *all* the states of affairs. This, then, must be the fact or state of affairs that the great conjunction *is* all the states of affairs.

Armstrong 1997: 198

Totality facts are higher-order facts. More precisely, the totalling relation holds between a mereological aggregate and a property that *all* things that compose the aggregate have: 'A certain mereological object totals a certain property. The mereological object is the whole composed of the existents in question' (Armstrong 2004: 72). Such a fact has the form Tot(aggregate, the corresponding property) (Armstrong 2004: 73). For instance, if all humans are mortal, then the totalling relation holds between the mereological sum of all humans and the property *mortality*.

Armstrong's totalling facts are rather different from the general facts of Russell and Hossack. Armstrong's totality relation is a higher-order relation relating lower-order facts or a relation between a mereological sum and a property. Russell's and Hossack's general facts are not of this kind. Armstrong's totality facts do not have the universal *generality* as constituent. Armstrong's totality facts are relational facts: the totalling relation holds between a property and an object. The universal *generality* requires only a property to form a fact. Armstrong seems to think that these cases boil down to the same thing:

The centrally important case, though, is Russell's own case: the general fact that all the facts (states of affairs) of lower order are all such facts. The analysis in terms of the Tot relation is the same. The object is the mereological whole (or manifold) of all the lower-order states of affairs. The property involved in the relation is the very abstract one of *being a state of affairs*.

Armstrong 2004: 74

But this is not so. Armstrong's totalling relation is a many-place higher-order relation. The universal *generality* is a one-place higher-order relation. The universal *generality* is a universal that is instantiated by first-order universals.⁹ These differences in form shouldn't be run together.

An analysis of universally quantified propositions in terms of the totalling relation falls foul of Russell's argument for the existence of general facts. Armstrong's totalling facts have the form $\text{Tot}(a, F)$, where Tot is the higher-order totalling relation, a is a mereological aggregate of particulars and F is a first-order property. From a proposition corresponding to a fact of this form and the proposition that b is one of the aggregate, it does not follow logically that b has the property. From the propositions (1) 'The totality relation relates the mereological sum of all humans and *mortality*' and (2) 'Socrates is human' it does not follow logically that Socrates is mortal. This is because (1) has the form $R(a, F)$, where R is a second-order property and F a first-order property, and hence it is an atomic proposition. Anyone taking Russell's argument for general facts from logical inferences seriously should be equally wary of analysing them in terms of Armstrong's totalling relation.¹⁰

'Totalling' is a more accurate name than 'alling' for Armstrong's relation, as totalling has little to do with 'all', that is universal quantification. Hossack's general facts are structurally isomorphic to universally quantified propositions, $(x)Fx$, where ' (x) ' refers to the higher-order universal *generality* and ' F ' to a first-order property. Armstrong's are not. It is, therefore, preferable to follow Hossack and to analyse general facts as sharing a common universal *generality* rather than in terms of a totalling relation.

Hochberg makes some pertinent observations about how Russell's argument for the existence of general facts fares if there are infinitely many objects. In this case, it is no longer possible to write down a premise stating what every thing is identical to: this would require an infinite disjunction. Someone not worried about the finitary limitations of predicate logic may simply say that although there is no such sentence, there is nonetheless such a proposition and leave it there. We can, however, do a little better than that and use the infinite case to argue even more forcefully that there are general facts.

In the finite case, we were drawing on a complex property that every object in the domain has, namely being either a_1 or being a_2 or being a_3 or ... being a_n , a property to which we can refer using identity, disjunction and names for each object. We are looking for a similar property that every object has in case the domain is infinite. Let's suppose for simplicity that there are ω objects and that the language has a name for each of them, say a_i , for $i \in \omega$. One option of a property of every object in the domain is 'being identical to something'. This can be expressed using identity and existential quantification. But $x(x=a_i)$ is logically true in classical logic, so adding the infinite collection of premises $x(x=a_i)$, for all i , to an infinite collection of premises Fa_i , for all i , does not get us any further towards a valid inference of $(x)Fx$ than not adding them. Using a free logic, where $x(x=a_i)$ is a substantial claim, makes no difference to the issue at hand. We need a premise to the effect that *nothing else* but the a_i exist, and even an infinite list of premises $x(x=a_i)$ does not give us that.¹¹

We would require a premise that collects together all the things there are, so that we can conclude that an arbitrary object referred to by a free variable, say x , is one of them, and then something that allows us to infer that x must be identical to one of the things named in the premises Fa_i . Alternatively, assuming $\neg Fx$, where x is a fresh free variable, we must be in a position to derive that x is one of all the things there are, so that we reach a contradiction, as for any thing there is, we have a premise to the effect that it is F . So not only do we need premises stating the F ness of each object and a premises stating that some property is had by all the things there are; we also require a premise connecting the infinitely many premises Fa_i , for all i , with the premise stating a property every a_i has. In other words, we need an additional premise that all the things there are are F s, and we have gone nowhere in the attempt to avoid general statements of the form $(x)Fx$, which correspond to general facts. To conclude with Hochberg, in the infinite case 'every true generalisation would, in fact, require its own fact' (Hochberg 1969: 341).¹² The infinite case, so far from being a problem for Russell, is grist for the mill of Russell and Hossack: it shows even more forcefully that there must be general facts. And this is particularly so for Hossack, for whom general facts consist of universals, such as *being an F*, instantiating the universal *generality*.

To close this section, a few words on whether general facts are positive or negative may be in order. Metaphysicians are in two minds over this question. For Armstrong, they are negative; for Hochberg, they are positive. For Hossack, too, general facts look rather more like positive than negative facts. Negative facts have as a constituent the universal *negation* in principal position. General facts don't. According to Armstrong, we should avoid accepting negative facts in our ontology that are on a par with atomic facts. To do so, Armstrong argues that it is necessary to accept that there are totality facts that set limits, and as such they are negative. 'The *all* state of affairs is itself a "no more" state of affairs' (Armstrong 2004: 58, see also 70). It is a negative state of affairs, albeit a higher-order one. One kind of negative fact suffices to explain all first-order negative truths (Armstrong 1997: 200). The general facts Hochberg considers are essentially described by lists, and as such they are positive.

Another option is to follow Ayer. Ayer observes that a general distinction between negative and affirmative statements is difficult, if not impossible to draw; at least so far no one has provided a satisfactory one. It looks likely that 'the distinction between

affirmative and negative statements [is] a matter of emphasis' (Ayer 1952: 815). Some such distinction can, however, be drawn, even if it may be one with less of a metaphysical impact than one might expect. Positive statements are more specific than negative ones. 'This flower is not yellow' leaves open a range of colours the flower may have which is not left open by 'This flower is blue'. This is why negative statements cannot be reduced to positive ones, although the less specific negative statements are entailed by the more specific positive ones. Ayer concludes that 'logically a negative statement [...] can be verified only through the truth of some more specific statement which entails it; a statement which will itself, by contrast, be counted as affirmative' (*ibid.*). A general statement, being one that entails much specific information, should then count as rather more positive than negative.

Russell, on the other hand, rejects that the distinction between negative and positive is applicable to general facts, because 'all' is the negation of 'some not', 'some' the negation of 'all not': 'In regard to general propositions, the distinction of affirmative and negative is arbitrary [...]. All general propositions deny the existence of something or other' (Russell 1919b: Lecture V, 190f). In light of these disagreements, it may be, as it is so often, wise to side with Russell.¹³

6. Conclusion

Russell concludes his discussion of general facts with an admission and an exhortation:

I do not profess to know what the right analysis of general facts is. It is an exceedingly difficult question, and one which I should very much like to see studied.

1919: Lecture V, 201

Hossack has taken up the challenge and proposed a bold thesis about the analysis of general facts, in keeping with his rationalist and realist outlook. Russell and Hossack agree that there are general facts, but Hossack is more forward concerning their nature. Following in their footsteps, I have proposed three arguments for Hossack's account of general facts and the existence of a universal *generality*. They are Russellian in spirit and draw heavily on Russell's writings.

Notes

- 1 My greatest debt in coming to grips with the topic of this paper is to Keith. It has also profited greatly from discussions with Javier Cumpa and a seminar in London with, besides Keith, Julien Dutant, Dorothy Edgington, Guy Longworth, Hamid Taieb and Mark Textor. Thank you also to Jonathan Nassim and Bahram Assadian for comments on the final version. While finishing the chapter I was supported by the Alexander von Humboldt Stiftung during a time as a research fellow at the University of Bochum.
- 2 Kürbis (ms.) is a historical investigation into Russell's changes of mind concerning general facts and different options of what their nature might be. I should mention a

third option regarding the latter issue. Elkind (forthcoming) argues that Russell had the material at hand for an account on which completely generalized propositions, those of formal logic, are made true by completely general facts, facts that have no constituents at all and which are non-atomic, non-molecular 'logical structures' (Elkind forthcoming: 21). On Hossack's account, completely general facts contain only logical universals.

- 3 This may raise the question whether there can be universals under which only one thing falls, such as *being identical to Socrates*. Then there is only a resemblance in the rather strained sense that Socrates resembles himself in this respect. Some aspects of resemblance may be explicable or definable in terms of others, so not all of them need to be assumed as unanalysed primitives. In the ultimate analysis, not every aspect of resemblance needs to correspond to a universal. For this reason, the examples of universals to be given in this chapter, always put in *italics*, are merely heuristic.
- 4 See for instance Hale (2013: ch. 1). See also my review of Hale's book (Kürbis 2015b).
- 5 The reason may have to do with Russell's avoidance of referents for logical symbols: whether these kinds of facts are interdefinable depends on whether, and if so, how, forms of facts can be embedded. Depending on the answers, an account that classifies facts according to their forms may require more primitive forms than an account that classifies them by principal and sub-principal universals requires primitive universals.
- 6 In correspondence Hossack suggested a further argument for accepting existential or general facts. The real resemblances are the natural divisions underlined by Plato, and these accord with the laws of nature. The laws of logic are laws of nature, and hence there must be logical universals. Instantiation and generality are essential to the laws of logic, and so they are both plausible candidates for logical universals. Whether to choose only one, and if so, which one, or both as primitive may depend on further considerations.
- 7 This reconstruction of Russell's argument follows Hochberg (Hochberg, 1969, 336f), about whom more later.
- 8 See also (Russell 1919a: Lecture III, 42).
- 9 This may have to assume that there are compound universals. But if some universals are related to others, this is hardly an issue.
- 10 This appropriates an argument of Bergmann's discussed by Hochberg (1981: 386). Hochberg reports Bergmann's argument to be one 'against the construal of a notion of "essential connection" in terms of a second-order relational property', i.e. against the claim that the notion of an essential connection between first-order properties F_1 and G_1 is to be analysed in terms of a second-order relation C_2 , such that $C_2(F_1, G_1)$: if there is an essential connection between F_1 and G_1 , then any F_1 is G_1 , but $C_2(F_1, G_1)$ does not logically entail $(x)(F_1x \supset G_1x)$, as $C_2(F_1, G_1)$ is an atomic proposition. $(x)(F_1x \supset G_1x)$ expresses, of course, a second-order relational property, but the crucial point is that it is not an unanalysed relation, but one formed from F_1 and G_1 by means of logical constants.
- 11 Besides, in a negative free logic, probably the free logic preferred by metaphysicians of the Russellian type, if F is atomic, $Fa_i \vdash x(x = a_i)$, so once more nothing is added to the argument by adding these premises.

- 12 Hochberg's route to his conclusion is, however, flawed. Hochberg fails to distinguish questions concerning the number of things there are and questions concerning which things there are. He writes that 'we can maintain, in view of [Russell's argument], that there is only one general fact for a finite universe: that pertaining to how many things (at most) there are' (Hochberg 1969: 334). Concerning the infinite case, Hochberg writes that 'together with the sentence used to state that there were at most a denumerably infinite number of objects, no set of atomic (and negations of atomic) sentences would suffice as premises to derive a generality of the form " $(x)(Fx \supset Gx)$ "' (Hochberg 1969: 341). But this is true in the finite case, too. Even if I add 'There are at most two objects' as a premise to the argument that attempts to conclude $(x)Fx$ from Fa and Fb , unless I know that my premises name them both, rather than one of them twice, I'm none the wiser. The premise $(x)(x=a \vee x=b)$ cannot be replaced by the premise that there are at most two things in the domain, $(z)(y)(x)(z=x \vee z=y)$. The latter requires the further premise $a \neq b$.
- 13 For further discussion of negative facts and the logic, metaphysics and epistemology of negation, see my (2015a), (2018), (2019a), chapter 4 of (2019b), and (2019c). The last item is a commentary on McCabe (2019).

References

- Armstrong, D. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Armstrong, D. (2004). *Truth and Truth-Makers*. Cambridge: Cambridge University Press.
- Ayer, A. J. (1952). 'Negation'. *The Journal of Philosophy* 49(26): 797–815.
- Elkind, L. D. C. (forthcoming). 'An argument for completely general facts: Generalized molecular formulas in logical atomism'.
- Frege, G. (1884). *Die Grundlagen der Arithmetik*. Breslau: Verlag von M. & H. Marcus.
- Hale, B. (2013). *Necessary Beings*. Oxford: Oxford University Press.
- Hochberg, H. (1969). 'Negation and generality', *Noûs* 3(3): 325–343.
- Hochberg, H. (1981). 'Natural necessity and natural laws'. *Philosophy of Science* 48(3): 386–399.
- Hossack, K. (2007). *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Kürbis, N. (2015a). 'What is wrong with Classical Negation?' *Grazer Philosophische Studien* 92(1): 51–86.
- Kürbis, N. (2015b). Review of Bob Hale: *Necessary Beings*. *Disputatio* 7(40): 92–100.
- Kürbis, N. (2018). 'Molnar on Truthmakers for Negative Truths'. *Metaphysica* 19(2): 251–257.
- Kürbis, N. (2019a). 'An argument for minimal logic'. *Dialectica* 73(1–2): 31–62.
- Kürbis, N. (2019b). *Proof and Falsity. A Logical Investigation*. Cambridge: Cambridge University Press.
- Kürbis, N. (2019c). 'The importance of being erroneous'. *Australasian Philosophical Review* 3(2): 155–166.
- Kürbis, N. (ms.). 'Russell on generality 1910–1918'. *In preparation*.
- McCabe, MM. (2019). 'First chop your *logos* . . . : Socrates and the Sophists on language, logic and development'. *Australasian Philosophical Review* 3(2): 131–150.
- Plato (1997). *Complete Works*. Edited by J. M. Cooper and D. S. Hutchinson. Indianapolis, IN: Hackett Publishing Company.

- Russell, B. (1911). 'On the relations between universals and particulars.' *Proceedings of the Aristotelian Society*, New Series, vol. 12 (1911–1912): 1–24.
- Russell, B. (1912). *The Problems of Philosophy*. New York: Henry Holt and Company; London: Williams and Norgate.
- Russell, B. (1918). 'The philosophy of logical atomism [Lectures I and II]'. *The Monist* 28(4): 495–527.
- Russell, B. (1919a). 'The philosophy of logical atomism [Lectures III and IV]'. *The Monist* 29(1): 32–63.
- Russell, B. (1919b). 'The philosophy of logical atomism [Lectures V and VI]'. *The Monist* 29(2): 190–222.
- Russell, B. (1919c). 'On propositions: What they are and how they mean.' *Proceedings of the Aristotelian Society*, Supplementary Volumes 2: 1–43.
- Russell, B. and Whitehead, A. N. (1910). *Principia Mathematica*. Vol. I. Cambridge: Cambridge University Press.
- Russell, B. and Whitehead, A. N. (1997). *Principia Mathematica to *56*. Cambridge: Cambridge University Press.

We Belong Together: A Plea for Modesty in Modal Plural Logic¹

Simon Hewitt

‘We belong together,’ a doe-eyed lover says to his beloved. She, being both better educated in a certain sort of metaphysics than her suitor and more romantic, responds, ‘Yes. Not only in the actual world, but in every possible world. There is no possible world in which we are not together.’ Whether this is genuinely touching or simply emetic is a fine judgement. Perhaps more remarkably, this expression of lovestruck hyperbole bears a striking similarity to a principle enshrined as current orthodoxy in discussions of the modal logic of plurals. Let xx be some things; then, goes one half of the claim, if x is one of xx then necessarily: if xx exist, then x is one of xx . Similarly, if x is not one of xx then necessarily: if xx exist, then x is not one of xx . The path from love to logic is a short one indeed.

Why should this be of any more than esoteric interest? The principle, which sometimes gets called *plural rigidity*, is put to work to important and controversial metaphysical ends. Crucially, Williamson deploys it in arguing both for necessitism (the doctrine that ontology is modally invariant) and for a property-based interpretation of second-order quantification, since he thinks the modal behaviour of plurals rules out a Boolos-style plural interpretation. (See Williamson 2010, 2013 and Boolos 1998b.) Linnebo develops a foundational programme for set theory in a modal plural logic strengthened by the addition of statements equivalent to the principle, and whilst he himself proposes a non-standard interpretation of the modalities, its acceptability with respect to metaphysical modality would provide a fall-back position for someone sympathetic to the approach (Linnebo 2013). On the other hand, the project of providing a plural paraphrase of talk about complex objects, explored by Hossack in his (2000), might seem to sit uncomfortably with the principle.

A number of current philosophical projects, then, turn on this particular claim about the modal behaviour of pluralities.² Yet it is often adopted swiftly without the sustained argument which might reasonably be required to sustain a claim with such wide-ranging consequences.³ Recently, Linnebo has moved in the direction of making good this deficiency (2016). The present chapter argues that Linnebo fails to supply sufficient reasons to adopt the claim at issue, and that in the absence of alternative arguments a question mark must hang over standard appeals to strong modal logics of

plurals by philosophers. Of necessity, given the current state of the literature and the importance of Linnebo's work, much of what follows consists of a close criticism of Linnebo, but I go on to suggest that progress in the debate requires us to get clear about the relationship between language and non-linguistic reality and to pay close attention to linguistic use. These thoughts look likely to generalize across the philosophy of modal logic and to call into question much of the present fashion for modalizing.

The chapter proceeds as follows: §1 lays out the issue to be discussed, and undermines initial motivation for it; §§2–4 work through Linnebo's formal arguments, diagnosing a common justificatory circularity; §5 situates the sceptical position thus arrived at methodologically. I urge caution in the invocation of such a strong modal claim for philosophical argument in the absence of more persuasive arguments for its truth. I suggest that more attention needs to be paid to the use of language in making assertions and performing deductions as a key, and perhaps the only, constraint on deciding the acceptability of candidate logical principles. This position looks likely to have implications beyond the present case, and to call into question standardly accepted modal axioms.

1.

What Linnebo terms *plural rigidity*, we will call *plural necessitism*. The expression is still not perfect (it might suggest a deeper association with Williamson's metaphysics than can be supported), but at least avoids the danger of use-mention confusion that attaches to using the word 'rigidity' of a principle stated in the material, rather than formal, mode. Plural necessitism is equivalent to the conjunction of:

$$\begin{array}{ll} \text{(NecInc)} & x < yy \rightarrow \Box (Ex \wedge Eyy \rightarrow x < yy) \\ \text{(NecNInc)} & x \leq yy \rightarrow \Box (Ex \wedge Eyy \rightarrow x \leq yy) \end{array}$$

Here '*E*' is to be read as an existence predicate,⁴ allowing us to work within a variable domain framework. It is routine to modify (NecInc) and (NecNInc) for a fixed domain framework simply by removing the second conditional along with its antecedent conjunction in each formula. In what follows, versions of the principles with and without existence predicates will be denoted by '(NecInc)' and '(NecNInc)', with context serving to disambiguate.

As we have seen, plural necessitism plays a non-negligible part in motivating or supporting metaphysically contentious positions. What reason, then, is there to believe that it is true? It is provably independent of any normal modalization of **PFO**+ (Hewitt 2012). Motivation will therefore have to take the form of formal argument from independently compelling premises or of informal arguments.

As advertised above, our engagement will be mainly with the formal arguments Linnebo offers in support of plural necessitism (hereafter *pn*). Before that, some comment should be made on a case for the plausibility of *pn* which Linnebo develops by analogy with an argument for the modal invariance of set membership, based on extensionality and the necessity of identity.⁵ This proceeds from the principle:

$$(\text{Indisc}) \quad xx \equiv yy \rightarrow (\phi(xx) \leftrightarrow \phi(yy))$$

The left-hand side of the conditional abbreviates the claim that xx and yy are co-extensional. From (Indisc) is derivable:⁶

$$(\text{Cov}) \quad xx \equiv yy \rightarrow \Box(xx \equiv yy)$$

‘(Cov)’ here stands for *covariance*, the thought being that actually coextensive pluralities covary modally. Now (Cov) does not entail pn , since it is compatible with pluralities ‘drifting’ modally. Suppose that, with respect to some possibility w_1 , xx just are yy . In order for (Cov) to be satisfied, we require that at any possibility accessible relative to w_1 , xx just are yy . But now consider w_2 , accessible relative to w_1 , such that xx and yy both consist of the members of those pluralities with respect to w_1 with the addition of a , which does not exist with respect to w_1 . We have here no counterexample to (Cov), but pn fails.⁷

Linnebo takes this drift to be indicative of *soft extensionalism* about pluralities. This is the doctrine that extensionality holds within possibilities but not across them. He argues that this is an unstable position, for the same reason that he holds the parallel position regarding sets to be unsuccessful. With respect to the latter, he draws a comparison between sets and *groups*. We might think of groups as akin to social entities, such as the Supreme Court Justices or the Petrograd Soviet:⁸

The only reason to accept the principle of extensionality [for sets] ... is that a set, unlike a group, is fully specified by its elements. Thus, when tracking a set across possible worlds, there is nothing other than the elements to go on. This ensures that the tracking is rigid. By contrast, when tracking a group, there is more than the members to go on These considerations give rise to a dilemma that applies not only to sets but to any other notion of collection: either we have to give up the principle of extensionality, or else we have to accept the rigidity principle. There is no stable middle ground. Kripke famously taught us that there can be no ‘soft identity theory’ in the philosophy of mind, according to which mental states are identical with physical ones, but only contingently so, only the ‘hard identity theory’ committed to necessary identity. Our present conclusion is analogous. There can be no ‘soft extensionalism’ concerning sets of other kinds of collection, only ‘hard extensionalism’ that incorporates the rigidity claims and the idea of transworld extensionality that they embody.

Linnebo 2016: 459

This is surely correct as regards sets: as Boolos once put the point, if anything deserves to be called analytic of our concept of set, it is the axiom of extensionality (1998a), and the move from this recognition to the claim that possession of all and only its actual elements is an essential property of any given set is a natural one. The present issue, however, is whether this form of reasoning transfers across to the plural case. And here a number of questions arise. Most fundamentally, it is not entirely obvious that pluralities need be regarded as extensional in any sense: the temptation to think

this as immediate as it is in the set-theoretic case arises from the truism ‘a plurality just is its members’. So far, so uncontroversial: the cutlery represents no addition of being over and above the knife, the fork and the spoon. Does extensionality follow? It does not. Suppose that a plurality just is some things as picked out by an expression with a particular sense;⁹ then extensionality will fail. This would, amongst other things, open the door to the denial of the plurality/group distinction,¹⁰ a move which, in addition to ontological economy,¹¹ has in its favour the undoubted fact that many groups are picked out by noun-phrases (NPs) linguistically indiscernible from plural NPs. At this point questions would likely be raised about whether commitment to pluralities is any less costly than commitment to objects (Florio and Linnebo 2016).

Let’s grant extensionality for pluralities, at least for the sake of argument: problems remain with the argument. It is taken for granted that extensionality is captured by (Indisc), where modal operators may occur within the substitutends for the sentential meta-variables. Should somebody not antecedently disposed to accept *pn* accept this characterization of plural extensionality? Here is one reason why they ought not to: a non-singularist meaning theory for plural vocabulary could hold, consistently with the rejection of a parallel view for singular terms, that our referential practices with plurals are fine-grained, allowing the pluralities we track with them to come apart modally. After all, they might insist, openness to this kind of proposal is part of what it is to take plurals seriously as a linguistic phenomenon semantically distinct from singular vocabulary. And this being so, allowing modal operators to be substituted within (Indisc) is question-begging.

None of this is decisive either way. Sceptical doubt has been introduced to Linnebo’s plausibility argument, which ought perhaps to cause us to question whether we can formulate a logico-metaphysical account of pluralities independent of developing a better understanding of the function of plural expressions in language. We’ll return to that point in due course. Perhaps, though, we can short-circuit that imperative for present purposes. Might it not be that formal arguments for *pn* are available, showing that the contested principle follows from commonly accepted (or at least plausible) rules governing plurals? Linnebo thinks so, and supplies three arguments to that effect. To these our attention now turns.

2.

Here, and in what follows, we will work in the logic **PFO+** enriched with a singular existence predicate. We will allow ourselves the free use of modal operators, and leave the syntax of these intuitive. Linnebo’s first argument for *pn* we will call *Uniform Adjudication* (UA). Versions of UA are supplied for **PFO+** with and without an existence predicate. Since nothing important turns on which argument is at issue, we will focus on that without the existence predicate for the sake of simplicity.

Add to the language of **PFO+** a dyadic plural term forming operator ‘+’, which takes a singular term in its first argument place and a plural term in its second. The intended interpretation of this operator is as denoting adjunction, the operation of ‘adjoining

one object to a plurality' (Linnebo 2016: 664). The following seems plausible as a principle governing adjunction:

$$(UniAdj) \quad \Box \forall x (x < xx + a \leftrightarrow x < xx \vee x = a)$$

Linnebo doesn't provide an argument for (UniAdj), but one may be easily be found. The principle expressed within the scope of the necessity operator is, after all plausibly understood as (implicitly) definitional of the adjunction operator, and as such analytic and therefore necessary. Given (UniAdj), the next stage in (UA) is to assume $a < xx$. This gives us $xx \equiv xx + a$ by (UniAdj). We are licensed to necessitate this by (Cov), yielding:

$$(1) \quad \Box (xx \equiv xx + a)$$

The next stage is to argue that (UniAdj) yields:

$$(2) \quad \Box (a < xx + a)$$

According to Linnebo (UniAdj) *entails* (2) within the scope of our assumption, but no proof is provided. Let's talk through how one might go. Instantiating the bound variable with a , application of **K** to the right-to-left direction of (UniAdj) delivers:

$$(3) \quad \Box (a < xx \vee a = a) \rightarrow \Box (xx + a)$$

The necessity of identity and *modus ponens* give us (2). Returning to (UA), with (2) under our belts, we can appeal to (1) and get $\Box (a < xx)$. We discharge the assumption and we are done.

The problem with (UA) as a case for (NecInc) – remember that nothing here turns on the presence of an existence predicate – is the invocation of (Cov) to necessitate (1). Recall that (Cov) is derived using (Indisc):

$$(Indisc) \quad xx \equiv yy \rightarrow (\phi(xx) \leftrightarrow \phi(yy))$$

In the context of an argument for a component of *pn* this looks problematic. In order to get the derivation of (Cov) we need modal vocabulary to be admissible in the substitutends for the sentential meta-variables. But it is doubtful than anyone who doubts *pn*, and so anyone with worries about (NecInc), should allow this. For this admission expresses formally the doctrine that pluralities are extensional in a strong sense. Above, we observed that the admission of modal vocabulary in instances of (Cov) begs questions about the constraints on the semantic behaviour of plurals in modal contexts. This is a defeating worry for the use of (Cov) in (UA). Someone not antecedently committed to *pn* ought only to admit substitutends in (Indisc) that do not contain modal operators.

This is not enough to derive (Cov), so (UA) is not available to the non-*pn* believer. But then, in particular, it is not available when arguing with a non-*pn* believer; which is to say it lacks the required suasive force.¹²

3.

The second argument is owing to Williamson (2010: 699–700), and termed by Linnebo *partial rigidification* (PR). Considering a fixed domain context first, this proceeds from an instance of plural comprehension:

$$(4) \quad yy(xx \equiv yy \wedge \forall x(x < yy \rightarrow \Box x < yy))$$

Here, yy is a partial rigidification of xx . Now assume $a < xx$ and consider yy , the partial rigidification of xx . This gives us $\Box a < yy$. Together with $\Box(xx \equiv yy)$, which we have by (Cov), the derivation of $\Box a < xx$ is immediate. Discharging the assumption gives us (NecInc).

Linnebo considers this persuasive, but of limited value, since the fixed domain assumption is unlikely to appeal to anyone not signed up to Williamsonian necessitism.¹³ I doubt even that (PR) is of use in arguing for pn even modulo a fixed domain, because like (UA) it appeals to (Cov), and as we have seen this looks question-begging. Things look unfavourable also when (PR) is adapted for variable domains. To do this, we use a different instance of plural comprehension.

We proceed in a similar fashion as for the fixed-domain case, this time concluding that $\Box(Eyy \rightarrow a < yy)$ on the assumption that $a < xx$. Appealing to (Cov) would get us $\Box(Exx \rightarrow a < xx)$, allowing us to discharge and get (NecInc) only given that $\Box(Exx \rightarrow Eyy)$, but as Linnebo himself acknowledges this is not something that somebody who rejects pn is likely to allow.

$$(5) \quad yy(xx \equiv yy \wedge \forall x[x < yy \rightarrow \Box(Eyy \rightarrow \Box x < yy)])$$

If one is prepared to grant that yy are actually all and only xx , but contingently so, what stands in the way of granting the possibility that yy exist but xx do not?

For this reason, Linnebo takes (PR) to have force against the denier of pn in a fixed domain, but not in a variable domain context. He is certainly correct regarding the latter case, but it is unclear that the argument is any more convincing when the domain is fixed. As we saw, in that case (PR) still relies on (Cov), and at this point the pn -denier ought to refuse to follow. We arrive here at a theme recurring throughout the present chapter; formal arguments for pn are less successful than even Linnebo, a sober and balanced commentator, allows. The principle does not appear susceptible to proof from premises which are not too close to the principle itself for the proof to have suasive force. This will become more apparent when we consider the third and final of Linnebo's formal arguments.

4.

Say that a plurality is *traversable* just in case its members can be exhaustively listed. This is straightforward in the case of a finite plurality. For example, a three-membered plurality aa might be traversed as follows:

$$(6) \quad \forall x(x < aa \leftrightarrow x = a \vee x = b \vee x = c)$$

Linnebo says that we can assert *uniform traversability*, by which is meant the necessitation of wffs such as (6), thus, in the context of an argument for *pn* this might look fatally question-begging, a point to which we'll return in due course. First we need to deal with the case of infinite pluralities.

$$(7) \quad \Box \forall x(x < xx \leftrightarrow x = a \vee x = b \leftrightarrow x = c)$$

Here substantial infinitary resources are required.¹⁴ Working in plural $L^{\infty\omega}$, and adopting the notation of Hewitt (2012), uniform traversability in this case may be stated with *i* ranging over sufficient ordinals to index the members of *xx*,

$$(NDIS) \quad \Box(x < xx \leftrightarrow \bigvee_{\forall i} (x = a_i))$$

Given **T** in the background modal logic we can prove¹⁵ $\Box y < xx$ on the assumption that $y < xx$, giving us (NecInc) for fixed domains. A similar argument is forthcoming in when an existence predicate is predicate, giving us (NecInc) simpliciter. However, as Linnebo himself admits, '[the] premise of universal traversability is little other than an infinitary restatement of our target claim that a plurality is fixed in its membership as we shift our attention from one possible world to another' (2016: 667).

As advertised, there is little reason to believe that an assertion of uniform traversability will do anything to convince anyone not antecedently given to assent to *pn*. This much Linnebo himself concedes. However, the other formal arguments he deploys are in equally bad shape, as we have seen. In each case the complaint is the same: the argument will not persuade somebody not already disposed to assent to *pn* since at a crucial stage it deploys an argument that such a person has no reason to accept. We know that (NecInc), and therefore *pn*, is independent of any normal modalization of **PFO**+, so to the extent that these logics capture pre-formal reasoning about plurals and modality our target doesn't fall out of principles governing thought. What is more, *pn* has been questioned by a minority of philosophers. Still, my impression is that the great majority assent to it. Given the use to which the principle is put in philosophy, it would be unfortunate if it lacked adequate justification; yet the position we have reached suggests that this might be a live possibility. Can more be said?

5.

We might be tempted to think that too much is being conceded to scepticism. After all, a plural variable '*xx*' is supposed to formalize a natural language expression such as 'these things'. But, one can imagine an exasperated interlocutor protesting, if *this* thing is one of *these* things then of course it is necessarily the case that this is one of these things, or else they wouldn't be *these* things.

Unfortunately, this is simply an affirmation of faith in place of an argument: the claim that 'these things' designates all and only these things with respect to a

counterfactual situation is implicit in 'or else they wouldn't be these things.' The expression 'these things' is being used to justify the claim that of necessity these things are all and only these things. Yet it can only do this on assumption that of necessity 'these things' is semantically rigid in the sense that¹⁶ if it designates *anything* with respect to a counterfactual situation then it designates all and only the things it actually designates. This semantic presupposition is not *pn* – to identify the two would be a gratuitous confusion of use and mention; however it is clearly closely related. As we shall now see.

The formal logic of plurals is of interest because it allows us to codify plural concepts we express in natural language and to capture the rules of deduction governing these. Were **PFO+** merely a mathematical system, understood formalistically, then there would be no barrier to importing any interaction with modal operators we care to mention in order to extend the system. For as long as our concern is mere symbolic games, one game is as good as another. But on the intended interpretation of **PFO+** as a *plural* logic, the meta-theoretic statement which appropriately formalizes the claim that $\alpha\alpha$ is semantically rigid just in case if it designates *anything* with respect to a counterfactual situation then it designates all and only the things it actually designates is supposed to model a claim that could be made about natural language plurals. And a meaning theory for a natural language should be statable in (the same) natural language itself, worries about the semantic paradoxes notwithstanding. But now it is clear that the move to the metalanguage imposes no rigidity on object language plurals that wasn't already implicit in the statement of the object language *pn* itself, for ultimately the semantic principle for the formal language should impose nothing not present in the natural language which is our object of investigation. In that language: 'these things,' used in the statement of a theory of meaning for the language, means *these things*. The constraint on a meaning theory that it licences the disquotational statement of the semantic value of lexical items entails that the appeal to natural language in the fashion proposed by the indignant objector of the previous paragraph gives no new suasive purchase.

Where does this leave us with respect to *pn*? If a widely accepted logic of plurals does not deliver *pn* as a theorem, no matter which (normal) modal logic it is combined with, and if obvious formal arguments departing from the accentuation of plural logic with metaphysico-semantic principles concerning plurals, what possible motivation could there be for adopting *pn*? A temptation at this point is to appeal to more robustly metaphysical considerations about the nature of pluralities. Provisional on metaphilosophical choices minuted below, this gives the impression of an admission of defeat under the guise of further enquiry. For whatever the nature of pluralities might be, it is surely something that we approach through our canonical means of latching on to them, that is through the use of plural language. How else can I explore what it is to be a plurality other than through attention to our plural talk and reasoning? Do not be misled by comparisons with the scientific study of entities of some particular class. I can explore the nature of water through using an electron microscope: my use, notoriously, will not settle the question. But 'plurality' is not a sortal of the same kind as 'water': it expresses a general, formal concept applicable across a wide range of collections of entities of all sorts and categories. In its formality it belongs with 'object',

rather than ‘positron’; and the prospects for a substantial enquiry into the nature of objects, as opposed to a logico-semantic making lucid of the place of objects in the structure of the world as approached through language, are not good.

We cannot step outside of language to study the general structure of the world in its metaphysical purity, whatever that would involve. Such is our logocentric predicament. What is it about the use of plural terms – ‘these students’, ‘the Channel Islands’, ‘Pixies’ (if we take the last to be semantically plural, and not a singular name for a group) – that licences the application of *pn*? This, it seems to me, is the only question that could give us a good reason to adopt the principle. And here there nothing in use that proponents of the principle have thus far rallied to its support, except with respect to the proper subplurality of plural terms consisting of compound names – ‘Marx and Engels’, ‘2 and 17’ and these are not the contested cases.¹⁷ Moreover, there is some evidence from usage against *pn*: ‘You should be careful what you are saying; Smith could have been one of those people’, to adapt an example of Dorothy Edgington’s.¹⁸ To complain here that ‘those people’ is associated with descriptive content (‘the people talking in that corner at this party’) is simply to shift the dialectic, and not in a manner that looks likely to favour *pn*. Are plural terms associated with some kind of intensional content, a sense distinct from their reference? What would decide this issue? A prime datum here seems to be precisely the kind of case Edgington proposes, and this supports not *pn* but its negation, since it provides a *prima facie* example of an instance for which the universal claim fails.

I do not take this to be decisive. Perhaps there is an argument to be had in favour of *pn* that answers this objection. It might be, moreover, that reasonable constraints on a meaning theory motivate acceptance of the contested principle: what happens, for example, if we formalize modal plural logic along natural deduction lines, imposing the familiar requirements of harmony and conservativity on rules for logical vocabulary and treating plural terms correspondingly in an inferentialist fashion?¹⁹ The point is merely that the work of motivating *pn* philosophically remains to be done, and that in the absence of such work the appeal to the principle in support of strong metaphysical theses is illegitimate.

The frustrated defender of *pn* has another avenue of response. Isn’t to seek to ascertain the truth of the matter as regards *pn*, as we have been doing, in isolation from the project of developing our best overall theory of the world to succumb to methodological error? If the benefits of adding *pn* to one’s theoretical arsenal outweigh the costs, for example in terms of sitting uncomfortably with current usage, then one should adopt the principle. This after all is how science proceeds, and metaphysics and logic are part of total science. Thus, for example, Williamson has stressed the methodological continuity between ‘modal logic as metaphysics’ and science (2013: 422–429).²⁰ This position is, to my mind, unconvincing. The disanalogies between metaphysics and the natural sciences are legion, most obviously in the failure of metaphysical theories to make empirically testable predictions. It might be retorted that metaphysical theories can be tested against intuitions, perhaps by means of thought experiments, these providing something analogous to the experiments of natural scientists.

The role of intuitions in philosophy has been subject to a good deal of critical scrutiny of late (Eklund 2017; Cappelen 2012), and the analogy between thought

experiments and experiments simpliciter has been questioned, rightly in my view (Hacker 2009). However, even within the general methodological framework in question there would seem to be particular concerns about appeals to intuition to support *pn*. For here, if we are not to confine our attention to imagined scenarios of the Edgington type (which hardly offer unambiguous support to *pn*) we will be arguing about intuitions and/or counterfactual judgements concerning either the truth of sentences expressed using English plural terminology ('In this situation, this would be one of these') or philosophical terms of art ('plurality'). In the first case, the above mooted accusation of circularity recurs; in the latter, the supposed role of intuition recedes into the background, and we are left simply with an *a priori* metaphysical disputation of the old school, with no obvious similarity to the natural sciences or any other uncontroversially truth-conducive enterprise. Either way, moreover, there is a real danger of dominant views amongst current metaphysicians on *pn* reinforcing themselves by appeal to the intuitions of the already-convinced.

If what I have suggested above is correct with respect to *pn* then it looks likely to generalize beyond the case of *pn* to other modal intuitions to which common and important appeal is made in philosophy. It is commonplace for **S5**, or at least **S4**, to be appealed to as the 'right' modal logic, and for metaphysical investigation to proceed with one of these as the background modal logic.

There are disimilarities with the case of *pn*: notably appeals to model theoretic semantics and to counterfactuals in support of **S5**. In both cases, I contend that circularity worries similar to those noted about arguments for *pn* arise. Establishing that will have to wait for a subsequent paper. What is clear, however, is that in the absence of these arguments support for strong modal logics will rest on judgements concerning iterated modalities that are at least as precarious as those concerning expressions of *pn* in natural language. Perhaps we should just be more modest in our claims about modality.

Notes

- 1 Thanks to audiences at the Second Workshop of the European Non-Categorical Thinking Network, Turin, and in Leeds, and to John Divers, Øystein Linnebo, and especially to Keith Hossack, for discussions of material relevant to this chapter. It was under Keith's supervision of my PhD at Birkbeck College that I first thought about plurals, and it is good to be able to contribute to a Festschrift for a model of both philosophical rigour and intellectual generosity. The research leading to this chapter was part of the ERC Nature of Representation project at the University of Leeds, grant number 312938.
- 2 Here, and throughout, 'plurality' is a convenient singularizing locution, as is standard in the literature: a plurality is *some things*.
- 3 An earlier attempt to remedy this situation is Uzquiano (2011).
- 4 Strictly speaking, we may need to be more careful. We have not formally specified a lexicon, and in particular have not settled the question whether predicates may be untyped, in the sense that they may have argument places for which either singular or plural terms are eligible. From the perspective of tractable formation rules, proof theory and model theory, it is likely to be easier to type predicates. If the lexicon is typed, then 'E' is doing illegitimate double work in (NecInc) and (NecNInc) as

formulated. We can get round this straightforwardly, by using superscripts to signify the type of existence predicates in the formal specification of the lexicon, and then omit superscripts in practice as a harmless convention.

- 5 The case for the necessity of identity rehearsed by Linnebo is that developed by Barcan Marcus in her 1961.
- 6 Proof sketch: (1) $xx \equiv yy$ (Ass for CP), (2) $\varphi(xx) \equiv \varphi(yy)$ ((1, Indisc)),
(3) $(xx \equiv xx) \leftrightarrow \Box (xx \equiv xx)$ (Thm), (4) $(xx \equiv yy) \leftrightarrow (yy \equiv xx)$ (3, 2), (5) $\Box (yy \equiv xx)$
(1, 4, MP), (6) $xx \equiv yy \rightarrow \Box (xx \equiv yy)$ (1–5, CP).
- 7 Compare the counter-model to (NecInc) provided (using a fixed domain semantics) in Hewitt (2012).
- 8 In private correspondence Linnebo has raised the possibility of extending groups to non-social cases of non-rigid collections. At this point the difference between his position and that supported in this chapter becomes a subtle one. Many of what I think are pluralities, he thinks are groups. I do not think that an ontological commitment is incurred to anything over and above the members of the group, and I think that the surface grammar of the plural expressions typically used to denote these collections faithfully reflects their semantically plural status.
- 9 *Objection:* If a plurality is some things as *picked out via a particular sense* then it is not just some things. *Reply on behalf of the position:* The objection fails to take account of how central questions of ontological commitment are to the standard debate about plurals. I am claiming plural expressions do not carry ontological commitments not already implicit in singular terms referring to the constituent members of pluralities. This is *prima facie* compatible with attributing additional semantic content to plural expressions. It is up to my opponent to show otherwise.
- 10 A group being a collection of people picked out by some role or feature, such as the Supreme Court justices. See Uzquiano (2004).
- 11 Although an alternative view, congruent with recent advocacy of easy ontology would concede the existence of groups, since identity conditions for groups can be supplied in terms of the corresponding pluralities, but deny their explanatory role. On easy ontology, see Thomasson (2015).
- 12 Similar considerations tell against the admission of modal vocabulary in the comprehension schema appealed to in Uzquiano (2011: 237).
- 13 On which, see Williamson (2013).
- 14 The language of $L_{\infty\omega}$ expands that of first-order logic with identity by admitting disjunctions and conjunctions of any transfinite length, whilst allowing only finite blocks of quantifiers. We assume, moreover, that the language is equipped with a proper class of individual constants.
- 15 For details see Hewitt (2012).
- 16 The term ‘(semantically) rigid’ used of a plural NP is dangerously ambiguous. My usage here has it that $\alpha\alpha$ is rigid iff (a) if x is one of the things designated by $\alpha\alpha$ then x is one of the things designated by $\alpha\alpha$ with respect to any world at which $\alpha\alpha$ designates some things and x exists; and (b) if x is not one of the things designated by $\alpha\alpha$ then x is not one of the things designated by $\alpha\alpha$ with respect to any world at which $\alpha\alpha$ designates some things and x exists. An alternative usage would have it that simply $\alpha\alpha$ is rigid iff $\alpha\alpha$ designates the same plurality with respect to any world at which $\alpha\alpha$ designates any plurality. My view is that the latter claim is utterly harmless, indeed uninformative, since we have no grasp of what pluralities are, what it is for some things to be considered as many, apart from our use of plural language. Pluralities just are what plural expressions designate.

- 17 In many cases, the reference of 'we' is anaphoric for that of a compound name. Unfortunately, this makes the cute example with which the present chapter began something of a ladder to be taken away once the dialectic gets going.
- 18 Cited in Rumfitt (2005: 120–121).
- 19 This would be moderately surprising, since the inferential rules would have to produce a distinct logic from **PFO+** or any of its subsystems, the modalizations of which (NecInc) is provably independent. To the best of my knowledge, however, no work has been done on natural deduction formulations of plural logic, although it should prove to be a routine exercise involving the adaptation of a formulation of second-order logic. My point in mentioning the inferentialist appeal to rules is not to convey an optimism that progress might be obtained by this route, but rather to indicate the kind of project a result from which would provide an acceptable resolution of the present debate.
- 20 More recently he has appealed directly to data from natural science in support of modal claims: specifically dynamic systems theory is rallied to make a case for the Barcan principle and its converse (2016). There is a lot to be said about this, but in any case, it is not clear how the move could be generalized to support plural necessitism.

References

- Barcan Marcus, R. (1961). 'Modalities and intensional languages'. *Synthese* 13(4): 303–322.
- Boolos, G. (1998a). *Logic, Logic and Logic*. Cambridge, MA: Harvard University Press.
- Boolos, G. (1998b). 'To be is to be a value of a variable (or to be some values of some variables)'. In *Logic, Logic and Logic* pp. 54–72.
- Cappelen, H. (2012). *Philosophy Without Intuitions*. Oxford: Oxford University Press.
- Eklund, M. (2017). 'Intuitions, conceptual engineering, and conceptual fixed points'. In C. Daly (ed.), *Palgrave Handbook of Philosophical Methods*. New York: Palgrave.
- Florio S. and Linnebo, Ø. (2016). 'On the innocence and determinacy of plural quantification'. *Noûs* 50(3): 565–583.
- Hacker, P. (2009). 'A philosopher of philosophy: A review of Timothy Williamson's *The Philosophy of Philosophy*'. *Philosophical Quarterly* 59(235): 337–348.
- Hewitt, S. (2012). 'Modalising plurals'. *Journal of Philosophical Logic* 41: 853–875.
- Hossack, K. (2000). 'Plurals and complexes'. *British Journal for the Philosophy of Science* 51: 411–443.
- Linnebo, Ø. (2013). 'The potential hierarchy of sets'. *Review of Symbolic Logic* 6(2): 205–229.
- Linnebo, Ø. (2016). 'Plurals and modals'. *Canadian Journal of Philosophy* 46: 654–676.
- Rumfitt, I. (2005). 'Plural terms: Another variety of reference?' In J. L. Bermudez (ed.), *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*. Oxford: Clarendon Press, pp. 84–123.
- Thomasson, A. (2015). *Ontology Made Easy*. Oxford: Oxford University Press.
- Uzquiano, G. (2004). 'The Supreme Court and the Supreme Court justices: A metaphysical puzzle'. *Noûs* 38(1): 135–153.
- Uzquiano, G. (2011). 'Plural quantification and modality'. *Proceedings of the Aristotelian Society* 111: 219–250.
- Williamson, T. (2010). 'Necessitism, contingentism and plural quantification'. *Mind* 119(475): 657–748.
- Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford, Oxford University Press.
- Williamson, T. (2016). 'Modal science'. *Canadian Journal of Philosophy* 46: 453–492.

Aristotelian Aspirations, Fregean Fears: Hossack on Numbers as Magnitudes

Øystein Linnebo

1. Introduction

Hossack's new book, *Knowledge and the Philosophy of Number* (2020), revives the ancient view that numbers are magnitudes and thus a special kind of property, which are instantiated in the physical world. In this way, Hossack makes a promising attempt to solve some stubborn metaphysical and epistemological problems about numbers.¹

The book considers some of the most important number systems, such as the natural numbers, the reals and the ordinals. To say that the numbers from each of these systems are magnitudes is to claim that the system is a linearly ordered family of properties that are instantiated by certain types of quantities. What are these quantities? For present purposes, the most important thing to note is the Aristotelian claim that 'the most distinctive mark of quantity is that equality and inequality are predicated of it' (1941a: 6a27). For example, some objects – or a 'plurality', as I will often put it, purely for ease of communication – are a quantity. Two such quantities are equal just in case they are equinumerous, in the usual sense that they can be put in one-to-one correspondence. We can now formulate a simplified form of one of the central theses of the book.

The Magnitude Thesis (simplified form)

A magnitude is a property that is shared by equivalent quantities.

For example, the natural number 2 is a property shared by all pairs.

The resulting conception of numbers has some attractive features, philosophically as well as mathematically. Since numbers are properties instantiated in the physical world, they are philosophically no more mysterious than any other properties. In particular, there is no mystery about their applicability or relevance to the physical world.² As for mathematical virtues, the conception offers a pleasingly unified framework for various kinds of numbers. As Aristotle observed, different types of quantity share some important algebraic features. These shared features entail, via the Magnitude Thesis, that all magnitudes have a shared algebraic structure – that of so-called *positive semigroups*,

to be explained shortly. Thus, we obtain a unified treatment of different kinds of numbers, as was anticipated by Aristotle.³ Different kinds of number are associated with different kinds of quantity, and additional structure on each of these kinds of quantity gives rise to additional structure on the associated numbers.

I applaud Hossack for attempting to revive some very promising ancient ideas that have received too little attention in recent decades. His discussion also brings to light some broadly Aristotelian insights which are important and valuable. I will argue, however, that Hossack is too hostile to Frege and too narrow in his development of the Aristotelian insights. There remains a strong pressure to generalize beyond the scope of Hossack's account. While the potential for generalizations is often a benefit, it can also expose problems. And indeed, when the needed generalizations are made, we reintroduce some of the problems that have haunted Frege and the neo-Fregeans. I conclude that these problems cannot be skirted but need to be addressed head-on.

The structure of this chapter is as follows. First, I present Hossack's analysis of quantities. As we will see, these have both a mereological structure and support a notion of equality, which set the agenda for the next two sections. Then, I turn to Hossack's analysis of magnitudes. I query how different his magnitudes really are from mathematical objects, especially as these are understood in the Fregean tradition. Finally, I show that the pressure to generalize introduces a threat of paradox. Overall, I seek a rapprochement between Hossack's approach and that of my recent book, Linnebo (2018), in part by identifying some points of agreement, and in part by suggesting some ways in which his view might benefit from moving in the direction of my own.

2. Mereology based on an addition operation

Hossack's notion of quantity is explicitly Aristotelian. On p. 35, he quotes with approval the following Aristotelian definition:

'Quantity' means that which is divisible into two or more constituent parts.

1941b: 1020a7

In virtue of having 'two or more constituent parts', quantities have a mereological, or part-whole, structure. This means that quantities, as Hossack understands them, have both a mereological structure and support a notion of equality. I will discuss these two types of structure in this section and the next, respectively.

Let us start by considering some examples of quantities. In addition to pluralities, mentioned above, Hossack provides two further canonical examples of quantities, which he calls *continua* and *series* (2020, sect. 3.1). Continua are the spatial regions occupied by stuffs, which in turn are the kinds of things to which mass nouns refer; for example, this milk, that beer, or all your gold. A series is some objects in a certain order. For example, a queue is a series. A series can thus be specified by means of an ordered list. Alice, Beth and Cassie are a series, and so are Tom, Dick and Harry. Here the order matters, unlike in the case of pluralities.⁴

Each of these types of quantity has a mereological structure. This is particularly clear in the case of pluralities and continua, which in fact share most of their mereological structure. We can talk about one plurality being part of another and about two pluralities overlapping or being disjoint.⁵ Likewise, this gold can be part of all your gold, which in turn may overlap all the gold in the UK. Even series have a mereological structure, where parthood is understood as the notion $\leq$ of being an initial series; for example, we have:

Alice, Beth $\leq$ Alice, Beth, Cassie

The mereological structure shared by pluralities and continua is a familiar one, namely that of Classical Extensional Mereology. This is the most widely studied mereological theory, often regarded as the default or classical account. But Hossack provides an alternative, less familiar axiomatization of this theory. Where the more usual approach uses parthood as its core notion, Hossack uses addition. He chooses this alternative axiomatization because he finds it easy to understand and ‘clearly a priori’ when interpreted as concerned with a system of quantities (2020: 44).

Before we can state his axiomatization, we need some explanations. Let us write ‘&’ for the operation of addition. For example, we have:

my part-time students & my full-time students = my students

We next observe that various other mereological notions can be defined in terms of this addition operation. First, x is *part of* y , written $x \leq y$, iff $x \& y = y$. Second, x is said to be a *proper part* of y iff $x \leq y$ but $y \not\leq x$. Third, x and y *overlap*, written $x \circ y$, iff they have a common part; if not, x and y are said to be *disjoint*, written $x \perp y$. Finally, x is an *upper bound* for some quantities iff each of these quantities is part of x ; and x is their *least upper bound* iff, whenever x is an upper bound, then $x \leq x$.

Hossack’s axiomatization of Classical Extensional Mereology consists of eight principles which, echoing Euclid, he calls ‘common axioms’.

- (A1) *Closure*. Given x and y , there exists a quantity w such that $w = x \& y$.
- (A2) *Idempotent law*. $x \& x = x$.
- (A3) *Commutative law*. $x \& y = y \& x$.
- (A4) *Associative law*. $x \& (y \& z) = (x \& y) \& z$.
- (A5) *Remainder*. If y is a proper part of x there always exists a quantity w disjoint from y such that $x = y \& w$.
- (A6) *Overlap*. If z overlaps $x \& y$, then z overlaps x or z overlaps y .
- (A7) *Universe*. There is a quantity of which every quantity is a part.
- (A8) *Least upper bound*. If $x \phi(x)$, then there is a least upper bound all x such that $\phi(x)$.⁶

As is well known, there are heated debates in metaphysics about the correct mereological structure of reality. Is this clay part of the statue, or vice versa, or perhaps both? Is there a sum of the clay and my bicycle? All these debates concern the

mereological structure of *individuals*, Hossack contends, not *quantities*. He wisely sets these debates aside in order to focus on the mereological structure of different types of quantity. His concern is thus solely with what he regards as a purely logical notion of parthood, where things are more clear cut. Indeed, Hossack claims that, when the mereological notions are interpreted as concerned with pluralities or continua, not individuals, the axioms stated above are uncontroversial and *a priori*.

I agree that the mereological structure of quantities is less problematic than that of individuals and that many of Hossack's axioms are plausible on the mentioned interpretations. I am unable to follow him all the way, however, as I have argued in print against (A7) on the plural interpretation.⁷ There is good reason, I argue, not to accept a universal plurality: a plurality has to be properly circumscribed in a way that the universe as a whole is not. Since this is a minority view, however, I set it aside for now and proceed on the assumption that Hossack's axioms are appropriate.

Pluralities and continua have further structure beyond that described by Classical Extensional Mereology. Say that x is *an atom* just in case x has no proper parts. Then, for pluralities, we want an axiom of atomicity, to the effect that every object has an atomic part:

(A9) *Atomicity*. $\forall x \ y (y \leq x \wedge \forall z (z \leq y \rightarrow z = y))$.

By contrast, continua are *gunky*, in the usual sense that every part contains a proper part. Thus, for continua, we add the following extra axiom:

(C9) *Divisibility*. $\forall x \ y \ y < x$.

Careful readers may have noticed that a single type of variables are used to range over individuals, pluralities, continua and perhaps more. This is not a slip but fully intentional. Following Aristotle, Hossack holds that "everything there is" divide into the predicables and the impredicables, that is, into properties and objects (2020: 22). Indeed, he goes even further, adopting a type-free language the variables of which range freely over properties, individuals, pluralities, continua and series: if we wish to speak about one category of things in particular, we can do so by simply restricting the variables to that category (43).

It is natural to worry that this untyped language will lead to paradoxes. For example, the property of not self-instantiating appears to lead straight to a property-theoretic version of Russell's paradox. In response, Hossack points out that his properties are scarce, not abundant, which means that there is no obvious pressure on him to accept the mentioned property or others that would lead to trouble (2020: ch. 2).

What types of quantities are there? As we have seen, Hossack provides three examples: pluralities, continua and series. He offers no systematic discussion, though, of whether there might be further examples. Why shouldn't there be? Various philosophers and logicians, including myself, have defended the idea of superplurals, that is, a doubly articulated form of reference that stands to plural reference the way this stands to ordinary singular reference. For example,

(1) My children, your children, and her children played against each other.

is naturally interpreted as involving superplural reference.⁸ There are other possibilities too. Why shouldn't there be a 'multiplural' form of reference that differs from ordinary plural reference, not by taking order into account, as in the case of serial reference, but by admitting multiple occurrences of the same object?⁹ I may, for example, state that the winners of the race in the past three years are Alice, Beth and Beth, thus conveying the fact that Alice won once and Beth twice. In short, there seem to be a variety of purely logical parthood relations and an associated formal mereology. A systematic theory of generalized parthood relations has been developed by Fine (2010).

I don't know to what extent Hossack would be willing to countenance other types of quantities beyond those that he discusses. But it is hard to see how he could block the proposed examples and generalizations just mentioned. Admittedly, he makes a point of the fact that natural language has 'devices for plural reference, mass reference and serial reference' (2020: 34). So he might attempt to block further generalizations by arguing that there is no basis for this natural language. But this type of argument would be doubly problematic. First, it is doubtful that natural language is so limited, as the above examples illustrate. Second, and more importantly, appealing to what is available in a natural language such as English would anyway be too parochial for present purposes. One cannot assess whether a kind of quantity exists as a legitimate object of mathematical investigation by appealing to the structure of English or other natural languages.

3. Equality

As Hossack is fond of pointing out, Aristotle holds that 'the most distinctive mark of quantity is that equality and inequality are predicated of it' (1941a: 6a27). For example, two pluralities can be equally numerous, and two masses, equally massive. To develop this idea, we need a notion of equality of quantities, which is distinct from identity. I will write this equivalence as $\sim$.¹⁰

What assumptions do we need concerning this equivalence? The basic idea is clear. We need it to be an equivalence relation, which 'agrees with' our notions of addition and subtraction. When spelling this out, Hossack's source of inspiration is Euclid's famous Common Notions:

1. Things which equal the same thing also equal one another.
2. If equals are added to equals, then the wholes are equal.
3. If equals are subtracted from equals, then the remainders are equal.
4. Things which coincide with one another equal one another.
5. The whole is greater than the part.

Seeking technical improvement, though, Hossack adopts some different axioms which entail, but go beyond, all the Common Notions except (4), which is not needed at this stage. These are his *Equality Axioms*:

- (E1) Quantities equal to the same quantity are equal to one another.
- (E2) If disjoint equals are added to equals, the wholes are equal. That is:

$$x \perp y \wedge x \perp y \wedge x \wedge y \rightarrow x \& y \quad x \& y$$

(E3) (Trichotomy) Two quantities are unequal if and only if either the first is equal to a proper part of the second or the second is equal to a proper part of the first.

(E4) No quantity is equal to any of its proper parts.¹¹

We can define a notion of a quantity being less than another by letting ' $x < y$ ' abbreviate ' x is equal to some proper part of y '. The trichotomy axiom can now be formalized thus: for all quantities x and y , exactly one of the following is true: (i) $x < y$; (ii) $x = y$; (iii) $y < x$.

The basic idea, we recall, is that $=$ should be an equivalence relation that 'agrees with' the other relations among quantities. The Equality Axioms ensure that. Hossack's Lemma 5.5 establishes that $=$ is an equivalence relation. Moreover, (E2) states that $=$ is a so-called congruence with respect to addition, provided that proper care is taken to ensure that appropriate quantities are disjoint; and the analogous claim can be proved concerning subtraction. We can also prove that $=$ is a congruence with respect to $<$.

I wish to end the section by resuming my case for further generalizations. First, consider pluralities with the equivalence of equinumerosity. Then (E4) states that no plurality can be equinumerous with a proper subplurality. This means that no plurality can be (Dedekind) infinite! Thus, Hossack's analysis rules out infinite pluralities and their cardinalities, which seem like perfectly good magnitudes.¹² Second, linear orders are no doubt important. But why insist that *all* quantities be thus ordered? There are plausible examples of quantities that are not linearly ordered. Consider angles, which are naturally understood as quantities and support a familiar notion of addition. But the equivalence of angles is naturally understood *modulo* 2π , with the result that the ordering of angles fails to be linear. Or consider oriented line segments. These too can (in the Euclidean case) be shown to support natural notions of addition, subtraction and equality.¹³ What is shared by two line segments under this equivalence corresponds to a vector. But the order defined on these vectors in terms of our notion of addition is not linear.

I conclude that there is substantial pressure to generalize further than Hossack does. And again, it is unclear how he could resist this pressure, should he wish to do so.

4. Magnitudes

A magnitude, we recall, was explained as a property that is shared by equivalent quantities (Hossack 2020: Introduction, sect. 5). Let us now be more precise. Let a *system of magnitudes* be a family of properties associated with one and the same relation of equality (Introduction and ch. 1). We can now formulate what is perhaps the most important thesis of Hossack's entire book.

Magnitudes Thesis

Whenever there is a standard of equality that satisfies the Common Notions, there is a corresponding system of magnitudes such that quantities are equal if and only if they instantiate the same magnitude of the system.

Hossack 2020: 5

Let the variable m range over the system of magnitudes associated with some fixed equivalence $\sim$, and write ' $x \eta m$ ' for the claim that x instantiates m . Then the Magnitude Thesis tell us that:

$$(MT) \quad m(x \eta m \wedge y \eta m) \leftrightarrow x \sim y$$

There is another way to express this thesis as well. Let ϕ map a quantity to its magnitude. (MT) can then be written as a Frege-style abstraction principle:

$$(AP) \quad \phi(x) = \phi(y) \leftrightarrow x \sim y$$

This reveals an important structural similarity between the Magnitude Thesis and Fregean abstraction. The significance of this structural similarity will be a central concern in what follows.

As Hossack explains, we can now give an elegant account of the abstract structure of magnitudes. We have studied the two-fold structure of quantities, provided by mereology and the equivalence relation. The Magnitude Thesis enables us to transfer much of this structure from the quantities to the magnitudes that these instantiate.

We can, for example, define an operation of addition on the magnitudes. To see this, consider two magnitudes m_1 and m_2 . Suppose it is possible to find two disjoint quantities x_1 and x_2 that instantiate these magnitudes. We wish to define the sum $m_1 + m_2$ as the magnitude of the sum $x_1 \& x_2$. For this definition to be permissible, however, the magnitude assigned to $m_1 + m_2$, namely $\phi(x_1 \& x_2)$, must be independent of our choice of disjoint quantities x_1 and x_2 to instantiate m_1 and m_2 , respectively. So consider some alternative such choice, x_1 and x_2 . Since (E2) states that $\sim$ is a congruence with respect to the operation of fusing disjoint quantities, we have $x_1 \& x_2 \sim x_1 \& x_2$. By the Magnitude Thesis, this ensures the desired independence:

$$\phi(x_1 \& x_2) = \phi(x_1 \& x_2)$$

Thus, our definition is permissible.

With the operation of addition of magnitudes defined, we can now proceed to investigate its properties. First, we prove that, for any disjoint quantities x_1 and x_2 , the magnitude of their sum is the sum of their magnitudes:

$$(1) \quad \phi(x_1 \& x_2) = \phi(x_1) + \phi(x_2)$$

This is only the beginning. Using the correspondence between quantities and magnitudes afforded by the Magnitude Thesis, Hossack derives some important algebraic properties of the magnitudes. In particular, we can show that addition of magnitudes obeys the associative law and a law known as 'restricted subtraction':

$$a + (b + c) = (a + b) + c$$

$$a \neq b \text{ if and only if there is an element } d \text{ such that either } b = a + d \text{ or } a = b + d.$$

In technical parlance, we have thus shown that any system of magnitudes forms a so-called ‘positive semigroup’. Hossack calls this important result *the homomorphism theorem*.¹⁴ This theorem highlights two attractive features of Hossack’s account, which were advertised already in Section 1 of this chapter. First, the theorem offers a pleasing unification of different types of numbers or magnitudes. Provided that a type of quantity satisfies the axioms of mereology and equality set out above, the system of magnitudes which the quantities instantiate is guaranteed to have a certain algebraic structure. Of course, if we consider quantities satisfying other axioms, then the corresponding magnitudes may satisfy different algebraic principles. Second, some attractive Aristotelian explanations become available. Magnitudes are no more mysterious than any other properties that are instantiated in the physical world. In particular, the algebraic structure of these magnitudes can be explained in terms of an analogous structure on the quantities whose magnitudes they are. The algebraic structure of the magnitudes is, as it were, ‘inherited’ from an analogous structure on the quantities.

We observed above that the Magnitude Thesis is structurally similar to a Fregean abstraction principle: two quantities are associated with the same magnitude just in case they are equivalent. While this structural similarity is undeniable, its philosophical significance is obviously up for debate. Here Hossack and I disagree. He regards his broadly Aristotelian approach as fundamentally different from Fregean abstraction, for at least two reasons. First, there is an important structural difference, namely that magnitudes come with a linear ordering, whereas Fregean abstraction works on a greater variety of equivalence relations. Second, the things on the left-hand side of (AP) belong to the category of properties, not objects, as Fregeans would have it. Magnitudes are therefore acceptable to nominalists, who deny that there are abstract objects: for magnitudes aren’t objects at all but properties instantiated by ordinary physical objects.

I deny that these alleged differences are very significant. First, how important is it that each system of magnitudes is linearly ordered? In the previous two sections, I have argued that there is substantial pressure to generalize Hossack’s account so as to take on board magnitudes that aren’t linearly ordered, such as angles and vectors. I now wish to add a more principled, philosophical point. In my opinion, the single most valuable aspect of Hossack’s approach is that it makes available some very attractive, broadly Aristotelian explanations. It explains how we can latch on to certain abstract features of reality, understood as properties that can be instantiated in the physical world. Moreover, the algebraic structure of these properties can be understood as ‘inherited’ from an analogous structure on the entities whose properties they are. We are now in a position to make an important observation. *These attractive explanations have nothing to do with the equivalence on the relevant entities giving rise to a linear order.* On the contrary, the Aristotelian explanations apply to shapes, directions, and so on, just as much as to magnitudes. Consider the case of directions. A direction can be understood as a property of lines, such that two lines have the same direction just in case they are parallel.¹⁵

$$d(l_1) = d(l_2) \leftrightarrow l_1 \parallel l_2$$

Moreover, relations between directions, such as orthogonality (symbolized as ' $\perp$ '), can be seen as 'inherited' from corresponding relations between the lines whose directions they are:

$$d(l_1) \perp d(l_2) \leftrightarrow l_1 \perp l_2$$

This shows that the attractive, broadly Aristotelian explanations have nothing to do with the properties in question being linearly ordered.

Second, how significant is it that magnitudes are properties rather than objects? Hossack attaches great significance to this distinction (see 2020: Introduction and ch. 2). I disagree. To explain why, recall that Hossack uses an untyped logic where a single sort of variables is used to range over things of all types or categories, including objects, quantities and properties. First-order variables are thus allowed to have magnitudes as their values, and magnitudes are permitted to figure as members of pluralities and sets. I believe this gives us all that Frege and his followers ever wanted when they defended the idea of numbers as objects. Frege's notion of an object is a broadly logical one, and the primary point of reifying numbers is to be able to treat them as objects for various logico-mathematical purposes. But the licence so to treat the numbers is granted to us anyway by Hossack's choice of an untyped logic.

To elaborate, consider what is often held up as a key advantage of the Fregean thesis that numbers are objects, namely that this thesis licenses his famous bootstrapping argument for the existence of infinity many natural numbers. Here is the idea. Suppose we have established the existence of the natural numbers from 0 up through n . Since numbers are objects, this means we can consider the plurality of 0, 1 and so on up through n .¹⁶ Since the number of this plurality is $n + 1$, we have now established the existence of one more number. For this argument to go through, however, there is no need for numbers to be objects in some robust metaphysical sense that some philosophers might find problematic. All it takes is that numbers can figure as members of the pluralities whose numbers we consider. And this is something that Hossack grants us anyway via his untyped logic. Thus, if Hossack's numbers fail to be objects, this would be in some purely metaphysical sense, which Fregeans probably never sought and certainly do not need.

All in all, I am inclined to regard Hossack's properties and objects as just two different kinds of Frege-objects. I also believe that the essence of his broadly Aristotelian metaphysics and epistemology of numbers can be co-opted by the Fregeans – and to a large extent is already part of their view, through their emphasis on numbers as reified cardinality properties and Frege's applicability constraint, which requires the applicability to counting to be part of the very nature of numbers, not only 'grafted on' from the outside.¹⁷

5. The bad company problem

Suppose we generalize, as I have been urging, and allow a wider class of equivalence relations to define magnitudes or, more generally, to figure in (MT) and its variant

(AP). Suppose further, following Hossack, that magnitudes and properties in general belong to the same logical type as objects. Then we encounter a well-known threat of paradox. For there are many instances of the resulting abstraction scheme (AP) that are inconsistent or otherwise unacceptable. A famous example is (a plural variant of) Frege's Basic Law V

$$(V) \quad \{xx\} = \{yy\} \leftrightarrow \forall z (z < xx \leftrightarrow z < yy)$$

which famously falls prey to Russell's paradox. This is the so-called *bad company problem*, which has received a great deal of attention in the Fregean tradition.

How should we respond to the problem? This is a huge question, which has generated a great deal of debate. I will end with some all too brief remarks about the options, including Hossack's and the one that I favour.

The neo-Fregeans take the lesson of the bad company problem to be that we need to restrict the kind of type-lowering involved in (AP) – where φ maps a property or a plurality to an object – to a limited range of instances, including the case of cardinality abstraction. How should the range of legitimate forms of abstraction be demarcated? Despite repeated attempts, there has been no satisfactory answer to this crucial question.¹⁸

In addition to this worrying lack of progress, Hossack has an independent reason to reject the neo-Fregean response. As we have seen, he is committed to treating magnitudes and any generalizations thereof as things alongside ordinary objects, all in the range of his single untyped kind of variables. This means that there are no types to be lowered in the first place. Although this is a radical view, I find myself broadly in agreement. In particular, Hossack is right, it seems to me, that considerations about expressibility give us reason to lift the type restrictions in favour of a single untyped kind of variable.¹⁹

How should Hossack respond to the bad company problem? It is tempting to think that the problem arises only as a result of the generalizations I have been urging and that it can accordingly be avoided simply by refusing to generalize. This temptation is illusory, however, since the threat of paradox arises already for the kinds of abstraction that Hossack considers. To see this, recall that Hossack accepts unrestricted plural comprehension: any formula defines a plurality, provided there is at least one object that satisfies the formula. What about the well-ordered analogue of pluralities, namely series? One would expect a defender of unrestricted plural comprehension to be sympathetic also to unrestricted comprehension for series; that is, if a two-place formula defines a well-ordering without repetitions, then the formula defines a series. But as Hossack is well aware, this form of comprehension would blow up his theory by allowing a version of the Burali-Forti paradox. By unrestricted series comprehensions, there would be a series of all the ordinals. By Hossack's theory, this series would define an ordinal, which would have to be greater than all the ordinals – including itself. This paradox is a special instance of the bad company problem that arises independently of the generalizations I have been urging.

Hossack's response is to impose restrictions on what series there can be; in particular, he denies that there can be a series of all ordinals. The only series there are, he argues, are those that are defined by *recursive* well-orderings. This response seems to me both

unacceptably restrictive and ultimately *ad hoc*. To substantiate the former charge, consider all the real numbers. By the axiom of choice, this uncountable plurality can be well ordered. Why can't there be serial reference to these numbers in that order? This form of reference makes just as much sense, it seems to me, as plural reference to that uncountable lot of numbers, which Hossack accepts.

The *ad hoc* character of Hossack's response to the Burali-Forti paradox emerges when we compare it with his response to other paradoxes of naive set theory. Hossack responds to Russell's paradox by appealing to NFU – a version of Quine's New Foundations adapted so as to accommodate urelements. This means that there is absolutely no connection between his responses to the two paradoxes. The restriction to recursive well-orderings, which forms the heart of his response to the Burali-Forti paradox, plays no role whatsoever in his response to Russell's paradox.

As we have seen, Hossack's approach to numbers and magnitudes is heavily inspired by Aristotle. This makes it interesting, I think, to observe that my own response to the bad company can be seen as developing yet another Aristotelian idea, namely that magnitudes and generalizations thereof are *dependent entities*, which need to be accounted for in a bottom-up manner. This is a central theme of my recent book Linnebo (2018). The existence of magnitudes and other abstracta is explained in terms of their instances; and their properties and relations are explained in terms of corresponding properties and relations among the instances, namely by developing the 'inheritance' ideas adumbrated in the previous section. This results in a hierarchical conception of numbers and other mathematical entities, where entities higher up in the hierarchy depend in various ways on ones lower down.

This hierarchical conception informs my view of what should be accepted as permissible plural and serial reference. If we are serious about the hierarchy, I argue, every form of reference must take place at some stage or other of the hierarchy. Thus, to refer plurally or serially to some things, all of these things need to be present at some stage or other. This in turn means that each plurality or series needs to be bounded by some stage or other. We thus arrive a uniform and moderately liberal view of both forms of reference, which contrasts with Hossack's unhappy combination of an extremely liberal form of plural reference and a severely restricted form of serial reference. This moderately liberal view can also be seen to guard against Russell's paradox and Burali-Forti's, again in a uniform way.

In sum, I have found that there is much to like about Hossack's neo-Aristotelian approach to numbers and magnitudes. But I have argued there is a push to generalize further, which introduces many of the problems discussed in the neo-Fregean tradition. Finally, I have suggested that these problems can be addressed by seeking inspiration from yet another Aristotelian idea, namely that numbers and other mathematical objects are dependent entities.

Notes

- 1 See also Peacocke (2019), especially chs. 2 and 5, for a closely related project, which, however, is less self-consciously Aristotelian than Hossack's.

- 2 By contrast, Field (1989, pp. 18ff) and Kitcher (1978) argue that numbers, understood as abstract objects, would be mysterious in these ways.
- 3 See his *Posterior Analytics* 74a18–25, quoted and discussed in sect. 5.1.
- 4 See also the *serial logic* developed by Hewitt (2012).
- 5 For a seminal article on plural logic, see Boolos (1984); for a survey, see Linnebo (2017). See also Hossack (2000), where the mereological structure of pluralities plays a central role.
- 6 I have slightly modified Hossack's formulation so as to avoid any reliance on Quinean 'virtual classes'.
- 7 See Linnebo (2010) and Florio and Linnebo (2020).
- 8 See Linnebo (2017, sect. 2.4) for an overview and further references.
- 9 For the *cognoscenti*: Multiplural reference would thus stand to the well understood notion of multiset as ordinary plural reference stands to sets.
- 10 Hossack uses $\approx$, but I would like to reserve this symbol for equinumerosity (or some restriction thereof), as is standard in much of the literature with which we will engage, and to use $\sim$ for the completely general notion of equivalence.
- 11 I have rephrased this axiom slightly for greater clarity.
- 12 In fact, (E4) is not only unduly restrictive but also potentially in conflict with the rest of Hossack's theory. As we saw in the previous section, he accepts a universal plurality, that is, a plurality of everything whatsoever. Moreover, as his discussion of the natural numbers in ch. 6 makes clear, each natural number is part of this universal plurality, which accordingly is an infinite plurality. One would therefore expect the universal plurality to have proper parts, or subpluralities, that are equivalent to itself, in violation of (E4). Hossack is saved from outright contradiction only by his unusually restrictive definition, in sect. 6.2, of the relation $\approx$ of cardinality equivalence, or 'tallying', between two pluralities.
 Even setting aside worries about this relation being too restrictive, a problem remains: his proof in sect. 6.3 that $\approx$ is an equivalence, in the sense of the equivalence axioms (E1)–(E4), is flawed. If the proof were correct, $\approx$ would be reflexive, by Lemma 5.5. But since only finite pluralities can tally, the universal plurality shows that $\approx$ isn't reflexive. To remedy this flaw, one option might be to abandon axiom (E4) and adopt a less restrictive notion of cardinality equivalence. Another option might be to redo the entire investigation in a way that replaces the work currently done by equivalence relations by so-called *partial* equivalence relations, that is, relations that are symmetric and transitive but not necessarily reflexive.
- 13 For example, each such line segment can be represented as (P_1, P_2) , where P_1 is its starting point and P_2 its end point. We define $(P_1, P_2) \& (Q_1, Q_2)$ as the segment (P_1, R) , obtained by first shifting the segment (Q_1, Q_2) such that its starting point coincides with P_2 and then letting R be the point to which its end point has been moved. (As is well known, this operation of 'parallel shift' is not available when the geometry is non-Euclidean.)
- 14 It is unfortunate that Hossack does not compare his homomorphism theorem with analogous theorems from the established field of measurement theory; see e.g. Suppes (1951), especially Metatheorem A, or Krantz et al. (1971), for a more authoritative treatment.
- 15 As usual, I ignore the slight awkwardness that Frege's 'directions' lack orientation. To capture the ordinary notion of direction, we need to consider oriented lines or line segments under the equivalence relation of 'co-orientation', defined as parallelism plus sameness of orientation.

- 16 How do we obtain the number 0 when numbers are understood as properties of pluralities, which presumably must have one or more members? Here we may follow Hossack's own lead and pick any non-number to serve as a proxy for 0 (see 2020: ch. 6).
- 17 See Wright (2000) for a discussion and partial defence of this constraint.
- 18 For this criticism, see Studd (2016) and (Linnebo 2018, sect. 3.2). For a recent survey of the options, see Cook and Linnebo (2018).
- 19 See the passage from Hossack's book quoted in Section 2 of this chapter (p. 192), as well as ch. 2 of Hossack 2020 in general.

References

- Aristotle (1941a). 'Categories', trans. E. M. Edghill, in R. McKeon (ed.), *The Basic Works of Aristotle*. New York: Random House, pp. 7–37.
- Aristotle (1941b). 'Metaphysics', trans W. D. Ross, in R. McKeon (ed.), *The Basic Works of Aristotle*. New York: Random House, pp. 689–926.
- Boolos, G. (1984). 'To be is to be a value of a variable (or to be some values of some variables)'. *Journal of Philosophy* 81(8): 430–449.
- Cook, R. T. and Linnebo, Ø. (2018). 'Cardinality and acceptable abstraction', *Notre Dame Journal of Formal Logic* 59(1): 61–74.
- Field, H. (1989). *Realism, Mathematics, and Modality*. Oxford: Blackwell.
- Fine, K. (2010). 'Towards a theory of part'. *Journal of Philosophy* 107(11): 559–589.
- Florio, S. and Linnebo, Ø. (2020). 'Critical plural logic'. *Philosophia Mathematica* 28(2): 172–203.
- Hewitt, S. (2012). 'The logic of finite order'. *Notre Dame Journal of Formal Logic* 53(3): 297–318.
- Hossack, K. (2000). 'Plurals and complexes'. *British Journal for the Philosophy of Science* 51(3): 411–443.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number: What Numbers Are and How They Are Known*. London: Bloomsbury.
- Kitcher, P. (1978). 'The plight of the Platonist'. *Noûs* 12: 119–136.
- Krantz, D., Luce, D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York: Academic Press.
- Linnebo, Ø. (2010). 'Pluralities and sets'. *Journal of Philosophy* 107(3): 144–164.
- Linnebo, Ø. (2017). 'Plural quantification'. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2017 edition.
- Linnebo, Ø. (2018). *Thin Objects: An Abstractionist Account*. Oxford: Oxford University Press.
- Peacocke, C. (2019). *The Primacy of Metaphysics*. Oxford: Oxford University Press.
- Studd, J. P. (2016). 'Abstraction reconceived'. *British Journal for the Philosophy of Science* 67(2): 579–615.
- Suppes, P. (1951). 'A set of independent axioms for extensive quantities'. *Portugaliae Mathematica* 10(4): 163–172.
- Wright, C. (2000). 'Neo-Fregean foundations for analysis: Some reflections'. *Notre Dame Journal of Formal Logic* 41(4): 317–334.

Mathematical Structures, Universals and Singular Terms¹

Bahram Assadian

1. Structures and universals

Mathematical structuralism is the thesis that mathematics is about structures. More precisely, a categorical mathematical theory – namely, a theory all of whose models or systems are pairwise isomorphic – is about the ‘abstract structure’ that its isomorphic systems ‘have in common’. For example, second-order arithmetic is about the abstract structure that all ω -sequences have in common. Thus, not only is there a system of the natural numbers, arranged in order of succession, there is also a further entity, which is the abstract structure of this system, and of any other system isomorphic to it. Similarly, second-order real analysis is about the abstract structure that all complete ordered fields have in common. This realist, ‘non-eliminativist’ conception of mathematical structuralism contrasts with the so-called ‘eliminativist’ conception, which does not leave any room for the existence of abstract structures. The eliminativist view paraphrases away statements about an abstract structure as quantifications over all the systems of the corresponding mathematical theory.²

The thesis that a mathematical structure is something that isomorphic systems of a mathematical theory ‘have in common’ or ‘share’ clearly suggests that a mathematical structure is akin to, and even has to be identified with, a universal. The conception of mathematical structures as universals has received a highly influential treatment in Stewart Shapiro’s (1997, 2006, 2008) *ante-rem* conception. According to this view, for each isomorphism class of systems there is a *structural universal* that is shared by the systems belonging to that isomorphism class, and the particular objects that have this universal in common are all and only the systems in that class.

Mathematical structures are, in this view, structural universals, and hence are ‘composed’ of other simpler universals. The notion of composition will be left open in this chapter. Furthermore, in being complex, such universals have other simpler universals as ‘parts’ or ‘constituents’. One of the problems concerning structural universals is whether any sense can be made of the thesis that one universal is a part or constituent of another. Again, the notion of part or constituent will be left unspecified here.³ I will use the neutral term ‘sub-universal’ to refer to simpler universals ‘composing’ each structural universal.

A mathematical structural universal has *positions* as its sub-universals. Positions are purely structural entities. They are akin to place-holders that could be, to use Shapiro's metaphor, 'occupied' by any entity belonging to the domains of the systems of a given isomorphism class. For example, the positions of the ante-rem structure of the natural numbers could be occupied by the Zermelo's finite ordinals or the von Neumann's finite ordinals.

Shapiro (1997: 83–84) adds a new twist to the story: there is a perspective – 'the positions-as-objects perspective' – from which we view the ante-rem structure as a system of the isomorphism class to which it corresponds. This special system is abstracted from the universal corresponding to the isomorphism class by considering it from the positions-as-objects perspective, in which the positions of the universal are construed as 'bona fide objects'; i.e. as the referents of mathematical singular terms whose semantic function is to refer to particular objects.

It seems then that the motivation for introducing the positions-as-objects perspective is fundamentally a semantic one:⁴ The system that is extracted from the structural universal is supposed to be the unique system to which we can fix the referents of the vocabulary of the background theory, and thereby secure determinate reference. For example, the expression 'the natural number 0' refers to a particular object: the first position in the ante-rem structure of the natural numbers. Similarly, 'the structure of the natural numbers' refers to a particular object: the unique special system with a domain of positions-as-objects, bearing purely structural relations to each other.

Natural as it may be, the above semantic motivation behind the positions-as-objects perspective gives rise to two problems for the thesis that an ante-rem structure (structure, hereafter) is what the non-eliminativist structuralist has been looking for as a candidate for *abstract* structures. In this chapter, I will discuss two arguments against the ante-rem view.⁵

The first is Geoffrey Hellman's (2001: 195–196; 2005: 546) permutation argument, which aims to undermine the uniqueness of structures: structures themselves, just like their instantiating systems, are subject to the permutation trick – in the sense that each structure produces distinct and yet isomorphic copies. And if there are distinct and equally acceptable structures, then the 'one-over-many' conception of structures – as types, universals, forms, etc. – has to be given up. As a result, the reference to the positions of structures will be every bit as indeterminate as reference to the elements of their corresponding systems; and hence, there will be no need for a commitment to structures: we could have lived with the referential indeterminacy within *systems* all along.⁶

The second argument is due to John Burgess (1999: 286–287). According to this argument, the ante-rem structuralist's structure is supposed to be a system, which is an ordered set, and so consists of a collection of positions-as-objects with distinguished relations on them. But abstract structures as what isomorphic systems have in common – what Burgess calls 'isomorphism types' – are abstracts with respect to the equivalence relation of isomorphism. The 'criterion of identity' for abstract structures and that of ante-rem structures are distinct: the former is based on the isomorphism of the systems they are structures of, while the latter is the identity of the underlying

domain of positions, and of the relations between them. Since the associated criteria of identity are distinct, ante-rem structures cannot be abstract structures.

In this chapter, I will try to resist the above two arguments. I address these challenges from a broadly neo-Fregean perspective, though there will be some important divergences too. In §2, following Linnebo and Pettigrew (2014, §5), I set the stage within an abstractionist framework, in which structures and their positions are abstracted in terms of Fregean abstraction principles. In §§3–4, I discuss the permutation objection, and argue that the source of this problem lies in the positions-as-objects perspective, in which structures are taken to be special systems, and hence are subject to the permutation trick. The permutation argument can thus be resisted if we abandon this perspective, and instead revise what is taken to be the semantic function of singular terms. This semantic revision is encapsulated in a thesis recently defended by Hale and Wright (2012), Hale and Linnebo (2020) and Hossack (2020, ch. 1 and ch. 4), according to which singular terms can refer to universals as well as objects. Thus, contrary to what Frege had recommended, it is not the case that *only* objects can be the referents of singular terms. This would allow the structuralist to use, for example, the singular term ‘the structure of the natural numbers’ to refer to a unique universal, and would thereby liberate them from positing a special *system* with a domain of positions-as-objects. Just as the singular term ‘the universal of *being horse*’ refers to a universal, so does ‘the structure of the natural numbers’; it refers to the universal *natural number structure*. We can permute the sub-universals, namely the positions, of this structural universal in a way that results in an isomorphic *system* of that structure. But this system itself is not the universal *natural number structure*. This, however, does not mean that by using the singular term ‘the structure of the natural numbers’ we cannot successfully refer to the unique universal, rather than a system obtained from it by permuting its sub-universals.

In §5, I address Burgess’s argument, which can best be viewed in the abstractionist framework of §2. Burgess’s argument poses a serious challenge for the structuralist who is committed to the positions-as-objects perspective, for it is on this picture that the criterion of identity for structures, being identified as systems, is codified in terms of the identity of positions and the relations they stand in. However, once we dispense with this perspective, there will be room for the identification of structures – taken as structural universals – with abstract structures or isomorphism types. All the same, independently of this rejoinder, Burgess’s argument crucially rests on the Fregean thesis that concepts with distinct criteria of identity cannot overlap. As Fine (2002: 46–54) and Linnebo (2018: 166–169) have convincingly argued, however, this Fregean thesis is not without its own problems either.

2. Structures and abstraction

Following Linnebo and Pettigrew (2014, §5), I set the stage for structuralism within a Frege-style abstractionist framework. In §68 of his *Die Grundlagen*, Frege proposes what is now commonly known as Hume’s Principle, which codifies the rule that equinumerous concepts should be associated with one and the same cardinal number:

Hume's Principle $\#F = \#G \leftrightarrow F \approx G$

where $\#$ is a singular term-forming operator, and $\approx$ stands for a one-to-one correspondence between the objects falling under F and those falling under G . Along similar lines, the structuralist abstracts on two (set-sized) systems α and β in accordance with the principle that isomorphic systems should be associated with one and the same structure i.e. the structure of α , denoted by $[\alpha]$, is identical to the structure of β if and only if α and β are isomorphic:

Structure Abstraction $[\alpha] = [\beta] \leftrightarrow \alpha \cong \beta$

where $\cong$ stands for isomorphism, and 'the structure of', denoted by $[\alpha]$, purports to denote an abstraction-operator, acting on isomorphic systems to give a structure as its value. Structure Abstraction thus captures the conception of structures as abstract structures or 'isomorphism types'; i.e. as entities that isomorphic systems have in common in virtue of their being isomorphic.

In addition to the abstracted structure, however, the structuralist needs to abstract the positions of structures from their corresponding elements belonging to the domain of the background systems. For there is more to structures than is stated in Structure Abstraction: structures are made up of positions or places. Let α and β be two isomorphic systems whose shared structure is $[\alpha]$, and the elements a and b belong to the domains of α and β . We say that a and b 'occupy' the same position in $[\alpha]$ – more precisely, the position of a in $[\alpha]$ is identical to the position of b in $[\alpha]$ – if and only if there is an isomorphism f from α to β , such that $f(a) = b$:

Position Abstraction $[a]_{[\alpha]} = [b]_{[\alpha]} \leftrightarrow f(f: \alpha \cong \beta \wedge f(a) = b)$

where $[a]$, denoted by 'the position of', is an abstraction operator from elements to positions. For example, since there is an isomorphism between the Zermelo's finite ordinals and the von Neumann's finite ordinals, we can show, by Position Abstraction, that, for example, $\{\{\emptyset\}\}$ of Zermelo and $\{\emptyset, \{\emptyset\}\}$ of von Neumann have the same position in the structure of the natural numbers. Position Abstraction thus allows us to specify the *domain* of structures in terms of the domains of their systems: given a system α with the domain D , the domain $[D]$ of its corresponding structure is the set of the positions occupied by the elements of D .

In the next step, we need to give an account of purely structural relations among positions. For instance, consider the relation that holds between the positions $[a]$ and $[b]$ in $[\alpha]$. The relation $[R]$ holds between $[a]$ and $[b]$ in $[\alpha]$ if and only if R holds between the corresponding elements a and b in α . More generally, let α be a system and R an n -ary relation on the domain D of α , such that $D = \{x_1, x_2, \dots, x_n\}$. The domain $[D]$ of $[\alpha]$ is: $\{[x_1], [x_2], \dots, [x_n]\}$, and $[R]$ holds between the elements of this set if and only if R holds between $x_1, x_2, \dots, x_n$.

The structuralist is committed to a further principle, according to which isomorphic structures are identical:

Structure Identity $[\alpha] \cong [\beta] \leftrightarrow [\alpha] = [\beta]$

which follows from two assumptions about structures, both of which are uncontroversial from the structuralist's perspective. The first is isomorphism between each structure and its system: (1) $[\alpha] \equiv \alpha$. The second is the thesis that if two systems are isomorphic, they have one and the same structure: (2) $\alpha \equiv \beta \rightarrow [\alpha] = [\beta]$. (1) and (2) entail that there are no two distinct isomorphic structures. Hence, Structure Identity.⁷

The thesis that isomorphic structures are identical seems to be the standard account of the identity-condition for structures.⁸ It has an important semantical consequence: in the structuralist's view, our putative reference to the elements of various isomorphic systems is indeterminate, since – as Benacerraf (1965) has argued – no isomorphic system is a better reference-candidate than any other. However, according to the structuralist, determinate reference to the positions of the structure that isomorphic systems have in common may nonetheless be achievable. According to her, referential indeterminacy would afflict positions of structures only if there were distinct and yet isomorphic structures, but Structure Identity is partly invoked to rule out this possibility, and thereby to secure the required uniqueness. It thus promises a way to accept a face-value construal of mathematical language, which contains a variety of proper names purporting to refer to objects. According to the structuralist, when we use '0', we are not talking about any initial element of an arbitrary system of arithmetic; we are rather referring to a particular object: the initial position in the structure of the natural numbers.

There will, consequently, be no need to paraphrase the language of arithmetic in order to provide alternative translations which are acceptable for eliminativist projects. But as we shall see in the next section, the sort of semantic indeterminacy revealed by the permutation argument aims to undermine the structuralist's hope for securing determinate reference.

3. The permutation argument

The permutation argument against structuralism, if successful, shows that structures, just like their instantiating systems, are subject to indeterminacy, in the sense that we can always permute the positions of a structure, and thereby produce a distinct and yet isomorphic 'copy'. As a result, there will be no non-arbitrary way to privilege one over any other as *the* structure of the associated mathematical theory, and therefore, there will be no non-arbitrary way to fix the reference of the terms of the theory to a unique structure. The argument has been formulated by Hellman as follows:

Suppose we *had* the *ante rem* structure for the natural numbers, call it $\langle N, \phi, 1 \rangle$, where ϕ is the privileged successor function and 1 the initial place. Obviously, there are indefinitely many other progressions, explicitly definable in terms of this one, which qualify equally well as referents for our numerals and are just as 'free from irrelevant features'; simply permute any (for simplicity, say finite) number of places, obtaining a system $\langle N, \phi, 1 \rangle$, made up of the same items but 'set in order' by an adjusted transformation, ϕ . Why should this not have been called 'the archetypical *ante rem* progression', or 'the result of Dedekind abstraction'? We

cannot say, for instance, 'because 1 is really *first*', since the very notion 'first' is relative to an ordering; relative to ϕ , 1, not 1, is 'first'.

Hellman, 2005: 546

But what is $\langle N, \phi, 1 \rangle$? Hellman calls it an 'ante rem structure', but it seems that he has, in effect, treated it from the positions-as-objects perspective, in which it is identified as a *system*, and thus any non-trivial permutation over its domain induces a distinct isomorphic copy. This is indeed where the ante-rem conception of structures succumbs to the problem posed by the permutation trick. For any system, albeit the special one abstracted from the corresponding universal, generates an isomorphic copy, and as Hellman correctly points out, this will be subject to the Benacerrafian embarrassment of riches.⁹

However, the problem with Hellman's argument lies exactly where it identifies each structure with a system of positions, and thereby presupposes that positions form a specific domain of objects. Hellman's argument shows that there are equally good systems. But it does not show the same for structural universals. Let $\langle N, \phi, 1 \rangle$ be a structural universal – namely, the universal that is shared by all the isomorphic natural number systems – and @ be a non-trivial permutation function over its positions or sub-universals. Let $\langle N, \phi, 1 \rangle^@$ be the result of this application. Since the permutation mapping is a set-theoretical operation, $\langle N, \phi, 1 \rangle^@$ is not a structure; it is a *system* instantiating the structure. The point is that we can permute the sub-universals of the structural universal *natural number structure* in a way that results in an isomorphic system, which is not itself the universal *natural number structure*.¹⁰

Thus, what has to be given up is exactly the positions-as-objects perspective. That is, we have to liberate the ante-rem conception from the thesis that each structure itself is a system, and therefore it is not a sort of an entity that can have distinct isomorphic copies. In fact, if we take structures as universals having their positions as sub-universals, the permutation argument loses its force.

As discussed above, however, Shapiro's motivation for introducing the positions-as-objects perspective is essentially a semantic one: to give a face-value reading of mathematical language, according to which what seems to be a singular term is indeed a singular term, whose semantic function is to refer to a particular object. Should we say that to give up the positions-as-objects perspective is thereby to give up the face-value construal of mathematical language? The answer would be in the positive only if singular terms were taken as expressions whose semantic function is to refer to particular *objects*. But this is not forced on us. If singular terms are taken to refer to particular entities, be they objects or universals, then we can still respect the face-value reading of mathematical language, without committing ourselves to the positions-as-objects perspective.

Thus, it seems that we can permute the sub-universals of the structural universal *natural number structure* in a way that results in an isomorphic system. But what does the singular term 'the structure of the natural numbers' refer to? It uniquely refers to the universal *natural number structure* and not to a system obtained from it by permuting its sub-universals. Since 'the structure of the natural numbers' is a singular term, and since it refers to the entity that all items of that type have in common, we can

take the expression to refer uniquely to the universal *natural number structure*. Here is an analogy: consider the first-level predicate ‘is a horse’, which is nominalized as the abstract noun ‘horse’. Let us suppose that when we use an abstract noun like ‘horse’, we can refer to the corresponding universal – that is, the universal *horse*. (I will come back to this use of ‘reference’ as standing for a relation between nominalized expressions and universals.) Thus, we manage to refer to a unique entity that all horses have in common. The referent is picked out by expressions such as ‘the universal *horse*’ or ‘the universal of *being horse*’. Likewise, when we use the nominalized expression ‘the structure of the natural numbers’, we refer to the universal that all the isomorphic natural number systems have in common.¹¹

According to the picture I proposed above, we can successfully refer to the universal *natural number structure* by using the singular term ‘the structure of the natural numbers’. Likewise, the singular term ‘the first position in the natural number structure’ refers to a sub-universal. This presupposes that we can refer to universals by using semantically singular terms. Thus construed, singular terms are expressions whose function is to refer to particular entities, universals included. I will discuss this thesis in the next section.¹²

4. Structures and singular terms

As noted above, Shapiro’s main motivation for introducing the positions-as-objects perspective is to account for a semantics in which expressions such as ‘the first position in the natural number structures’ are genuine singular terms referring to particular objects. This is indeed an inviting picture if the aim is to secure determinate singular reference to mathematical objects, despite the indeterminacy of reference afflicting the elements of the corresponding systems. After all, an important motivation for postulating structures, as a special sort of *system*, is to secure determinate reference. However, as we discussed above, this treatment of structures and their positions come at a serious price: they will be subject to the sort of indeterminacy exposed by the permutation argument.

In the previous section, I suggested that we can use expressions such as ‘the structure of the natural numbers’ to refer to the universal that all isomorphic natural number systems have in common, namely, the universal *natural number structure*. However, a universal is typically taken to be a kind of entity which cannot be *referred* to. This would presumably explain why Shapiro introduces the positions-as-objects perspective in which the referent of ‘the structure of the natural numbers’ is the special system of the natural numbers, as opposed to its structural universal. So, our question is this: Is there any prospect of using expressions such as ‘the structure of the natural numbers’ or ‘the second position in the natural number structure’ to refer to universals, without invoking the positions-as-objects perspective?

Instead of introducing the positions-as-objects perspective, the structuralist could revise the received account of the semantic function of linguistic expressions, according to which singular terms *only* refer to objects. Following Hale and Wright (2012), Hale and Linnebo (2020) and Hossack (2020, ch. 1 and ch. 4), the picture

I advocate here is as follows: a predicate such as ‘is the natural number structure’ stands for, or refers to, a universal. The predicate can be nominalized by means of nominalizations such as ‘the universal of *being natural number structure*’, ‘the universal that something instantiates in virtue of satisfying ‘is natural number structure’’, or simply ‘the natural number structure’. Then we can use these singular terms to refer to the universal expressed by the initial predicate. These nominalizations must not be taken to indicate that we use the corresponding singular terms to refer to particular *objects*. The key point is that singular terms can be used to refer to universals, concepts, properties and the like. Thus, instead of construing ‘the second position’ as a singular term referring to a particular object, we should say that it can be used to refer to the universal expressed by the predicate ‘is the second position’ by means of the relevant nominalization. Similarly, we can make singular reference to the universal expressed by ‘is horse’ by means of nominalizations such as ‘the universal of *being horse*’ or the abstract noun ‘horse’.

This requires a substantial modification of the Fregean doctrine that *only* objects can be the referents of singular terms, and that *only* predicates can refer to concepts and relations. According to the proposed modification of the Fregean doctrine, we can use singular terms to speak of entities of any ontological category. It seems perfectly natural and intuitive to speak of the property of *being wise*, the relation of *smaller than*, and so on. Once we have expressed the second-order existential generalization of ‘Mary and Bob are wise’ by ‘There is something (some universal) which Mary and Bob have in common’, we can hardly avoid answering questions such as “Which universal?” by ‘the universal of *being wise*’, or more simply: *wisdom*. It just seems obvious that in English and other languages employing higher-order resources that we refer to universals by means of singular terms of the type ‘the universal of being...’ or ‘the property of being...’.

Similarly, once we have expressed, for the systems α and β of arithmetic, the second-order existential generalization of ‘ α and β are isomorphic’ by ‘There is something (some universal) which α and β have in common’, we can hardly avoid answering questions such as ‘What?’ or ‘Which universal?’ by ‘the universal of *being natural number structure*’. Thus, on this picture, the structure of the natural numbers is just the universal referred to by means of the singular term ‘the structure of the natural numbers’. If this (neo-Fregean) account of reference to entities other than objects by singular terms is accepted, we can make singular reference to structures without any need to introduce a perspective in which they are taken to be systems, albeit special systems consisting of purely structural objects.

Interestingly, the abstraction principle Structure Abstraction manifests the sort of nominalization in play here: the term ‘the structure of the system α ’, symbolized by $[\alpha]$, is the nominalization of ‘being the universal instantiated by α ’. Similarly, Position Abstraction encodes a nominalization of predicates such as ‘being the second position in $[\alpha]$ ’ or ‘the universal being the second position in $[\alpha]$ ’, which are sub-universals of $[\alpha]$. These two abstraction principles, respectively, specify the conditions under which the reifications of structural universals and their positions are identical.

The nominalization operators apply to predicates to form new singular terms, which are the vehicles of singular reference to universals and, in particular, to

mathematical structures and their positions. Being singular terms, identity statements such as ‘the structure of the system α = the structure of the system β ’ are well-formed. The identity sign can be flanked by any singular terms involving nominalization operators. This generous conception of singular termhood ensures that ‘the structure of the system α = the structure of the system β ’ is a genuine identity statement, and yet the terms flanking the identity-sign are singular terms whose semantic function is to refer to universals.

To summarize: I have sketched a recent view of singular terms, according to which such expressions can be taken to refer to universals, and then applied it to mathematical structuralism. This generous conception of the semantic function of singular terms, if accepted, puts the structuralist in a position to abandon the positions-as-objects perspective, which is the source of the permutation problem. As a result, since structure-terms and positions-terms are taken as semantically singular, there will be a promising way for the structuralist to be faithful to the language of mathematics.

5. Ante-rem vs. abstract structures

In his review of Shapiro’s (1997), John Burgess argues that ante-rem structures cannot qualify as abstract structures: an abstract structure or an isomorphism-type is what isomorphic systems share, whereas an ante-rem structure is supposed to be an ordered set, and hence, consist of a collection of positions-as-objects with relations and functions on them: ‘Sometimes [ante-rem structures] are confused with isomorphism types, but this is a mistake: An isomorphism type is no more a special kind of system than a direction is a special kind of line’ (Burgess, 1999: 286–287).

However, the abstractionist framework proposed by Linnebo and Pettigrew, sketched out in §2, shows that in many cases, an isomorphism type can indeed be regarded as ‘a special kind of system’: Structure Abstraction encodes the information that the structure of α and the structure of β is what is shared by the isomorphism of the systems α and β . And Position Abstraction specifies a domain of positions-as-objects for the abstracted structures. However, as Linnebo and Pettigrew (2014: 276) argue, this treatment of structures is not available in all cases. In particular, where the background systems are ‘non-rigid’, i.e. admitting non-trivial automorphisms, Position Abstraction forces us to identify numerically distinct positions. Thus, in the end, Linnebo and Pettigrew uphold the negative thrust of Burgess’s remark.

In my view, however, Burgess’s argument can be resisted from a more general standpoint. But let us first see what his argument exactly is. On first pass, the argument is as follows: abstraction always takes F s to non- F s; e.g. lines to directions, concepts to cardinal numbers, systems to structures, and so on. In general, the equivalence-types are not to be identified with things they are types of. Thus, when we abstract on isomorphisms between systems to arrive at structures, we never get another structured system, let alone one isomorphic to the one we started with.

There is surely something right about this argument, but something is lacking, too. Why should we think that equivalence-types cannot be special sorts of things they are types of? There might still be a conceptual room for taking abstraction as a move from

*F*s to a special kind of *F*s. An important example is Cantor's abstractionist account of cardinal and ordinal numbers. In §1 of his *Beiträge*, Cantor defines a set as any collection *M* of definite and separate objects *m*: the elements of *M*. He then abstracts the cardinal number of any set as a special kind of set whose members are 'pure units', which have no distinguishing properties beyond their merely being distinct:

We will call by the name 'power' or 'cardinal number' of *M* the general concept which, by means of our active faculty of thought, arises from the set *M* when we make abstraction of the nature of its various elements *m* and of the order in which they are given. We denote the result of this double act of abstraction, the cardinal number or power of *M* [...]. Since every single element *m*, if we abstract from its nature, becomes a 'unit', the cardinal number [...] is a definite set composed of units, and this number has existence in our mind as an intellectual image or projection of the given set *M*.

Cantor 1955: 86

And in §7, Cantor gives a parallel account for the general concept of an ordinal number. His method for abstracting ordinals and cardinals is as follows: we abstract from the nature and any discerning properties of the elements of a given set to obtain its ordinal number, and then abstract from the order of the elements of the ordinal number to obtain its cardinal number. The key point for our purposes is that the cardinal number of a set itself is a set, albeit a special kind of set; i.e. one whose elements are pure units, having no distinguishing properties beyond their merely being distinct. Similarly, the ordinal number of a set itself is a set, albeit a special kind of set; i.e. one whose elements are pure units which are discerned only by the way in which they are ordered.¹³ It seems then that abstraction, as such, does not always move from *F*s to non-*F*s. In particular, as Cantor's account of abstraction shows, it can take *F*s to a special kind of *F*s.

There is, however, another argument that could lead to what underlies Burgess's claim, to the effect that abstraction always takes *F*s to non-*F*s. The argument is as follows: the criterion of identity for abstract structures is given in terms of isomorphisms of the systems they are types of; whereas the criterion of identity for ante-rem structures is given in terms of the identity of the underlying domain of positions, and of the distinguished relations on them. Hence, structures cannot be abstract structures.¹⁴ The argument can be formulated more precisely as follows: **P₁**: The criterion of identity for abstract structures is given in terms of the isomorphism of the systems they are types of. **P₂**: The criterion of identity for ante-rem structures is given in terms of the identity of the underlying domains and of the distribution of the relations over their positions. **P₃**: Where *x* is an entity of the sort *K₁* and *y* is an entity of the sort *K₂*, if *x* and *y* have different criteria of identity, then *K₁* and *K₂* cannot overlap. **C**: since the criterion of identity for abstract structures is different to that of ante-rem structures, they are different sorts of entities.

P₁ tells us what abstract structures as isomorphism-types are: they are things whose criterion of identity is specified in terms of the isomorphism of the systems they are types of. This is nicely captured in Structure Abstraction. **P₂** specifies the criterion of

identity for ante-rem structures, conceived of as special sorts of *systems*; i.e. the ones extracted from the structural universals corresponding to isomorphism classes by considering them from the positions-as-objects perspective. P_3 states the Fregean principle, according to which if two entities, x and y , receive distinct criteria of identity, the sorts to which they belong cannot overlap. P_3 encodes Frege's 'fundamental thought' about the inclusion and exclusion of concepts in terms of their associated criteria of identity. Following Hale and Wright (2001: 389), let us call the principle *Grundgedanke*: concepts with different criteria of identity cannot overlap. For example, given that the criterion of identity for cardinal numbers is specified in terms of one–one correspondence among concepts (as Hume's Principle makes it plain), and that the criterion of identity for persons is specified in terms of, say, psychological continuity, no number is a person. (This is an outline of the neo-Fregeans' solution to Frege's Caesar Problem.)

The *Grundgedanke* is an attractive principle. It accords well with the metaphysical role that criteria of identity are supposed to play: they give us information about the nature of the objects involved. For example, Hume's Principle is not only a stipulatively true implicit definition serving to fix the meaning of the numerical operator, it also tells us what numbers really are: they are *just* those things whose essence consists in facts concerning one–one correspondence among the associated concepts. Assuming that the criterion of identity for persons is different from that of the cardinal numbers, the *Grundgedanke* ensures that numbers and persons are essentially distinct entities.

Attractive as it may be, as Fine (2002: 46–54) and Linnebo (2018: 166–169) have convincingly shown, the *Grundgedanke* is not acceptable across the board. Consider the following two criteria of identity for cardinal numbers:

Hume's Principle	$\#F = \#G \leftrightarrow F \approx G$
Finite Hume's Principle	$\#_f F = \#_g G \leftrightarrow (F \approx G) \vee \neg \text{Fin}(F) \wedge \neg \text{Fin}(G)$

In the second principle, ' $\text{Fin}(F)$ ' abbreviates the second-order formulas asserting that F is finite. Thus, Finite Hume's Principle is just like Hume's Principle as far as finite concepts are concerned, but does not assign any number to infinite concepts.¹⁵ So, if F is finite, then the value of the numerical operators in Hume's Principle and Finite Hume's Principle will be the same: $\#F = \#_f F$. But if F is infinite, this does not hold.

Now, suppose that the criterion of identity for the concept *finite cardinal number* is given by Finite Hume's Principle, and that the criterion of identity for the concept *cardinal number* is given by Hume's Principle. Since these two principles incorporate distinct criteria of identity, then by the *Grundgedanke*, their corresponding concepts cannot overlap. But this clashes with the plausible claim that the finite cardinal number 2 and cardinal number 2 are identical. Therefore, either the finite cardinal number 2 is distinct from the cardinal number 2 or the *Grundgedanke* is false.

Accordingly, since P_3 in Burgess's argument rests on the *Grundgedanke*, it will be exposed to the problem we just canvassed. The *Grundgedanke* seems to fly in the face of some highly intuitive theses about the inclusion or exclusion of sortal concepts whose criterion of identity are in question, and hence, it cannot be universally accepted. Thus, it is consistent for concepts with different criteria of identity to overlap. As a

consequence, there seems to be a conceptual room for the concepts *abstract structure* and *ante-rem structure* to overlap, even though they receive distinct criteria of identity.

Furthermore, it appears that P_2 has its own problems, too. It should be already clear that in the picture I have been offering in the previous two sections, there is no room for the identification of an ante-rem structure as a special kind of system. But it is exactly this identification that provides the justification for P_2 . An ante-rem structure, as a structural universal, can be *represented* as a system; but it is not itself a system. Hence, contrary to P_2 , it is just not the case that an ante-rem structure is a sort of an entity whose criterion of identity is sensitive to the elements of its 'domain' and the distribution of relations over them. In sum, where there is no room for the positions-as-objects perspective, there will be no room for P_2 either.

Notes

- 1 I am very grateful to Øystein Linnebo, Jonathan Nassim and Richard Pettigrew for discussion and written comments, and also to John Burgess for a very helpful email correspondence about his argument, discussed in §5. Above all, I wish to record my debt to Keith Hossack, for his supervisions, encouragements and conversations ever since I started my graduate studies at Birkbeck College. Keith is no friend of structuralism; and I don't think I will ever succeed in convincing him about its attractions! This chapter invites him, once more, to see the merits of structuralism, as I believe that his philosophy of mathematics, and in particular, his conception of ordinals as universals, is consonant with structuralism in some important respects.
- 2 Following Shapiro (1997), I shall refer to 'models' and 'systems' interchangeably, as ordered $(n+1)$ -tuples consisting of a domain and relations on this domain.
- 3 For more discussion, see Armstrong (1986) and Lewis (1986).
- 4 See Shapiro (2008: 307, and 2006).
- 5 A terminological note: the term 'ante-rem' is due to Shapiro, with an explicit allusion to Platonic universals: *ante-rem* structures are ontologically independent from their instantiating systems. As will be clear below, the abstractionist framework for understanding structures enforces the idea that the identity-conditions of structures are explained in terms of their instantiations, which suggests an Aristotelian conception of structures as universals. In this chapter, the focus on *ante-rem* structures is directed on the thesis that they are universals, and that they are composed of positions. Nothing in what follows hinges on issues concerning ontological dependence. For more on this issue, see Linnebo (2008).
- 6 For more discussion on the permutation argument in the context of structuralism, see Assadian (2018).
- 7 The entitlement is straightforward. The right-to-left direction of Structure Identity is trivial. As for the other direction, assume $[\alpha] \equiv [\beta]$. By (1), $[\alpha] \equiv \alpha$ and $[\beta] \equiv \beta$. Hence, $\alpha \equiv \beta$, which, by (2), entails $[\alpha] = [\beta]$.
- 8 See Shapiro (1997: 93). Leitgeb (2021) identifies structures with unlabelled graphs, where the criterion of identity for unlabelled graphs G_1 and G_2 is this: $G_1 = G_2 \leftrightarrow G_1 \equiv G_2$.
- 9 Curiously, it seems that Shapiro endorses the conclusion of the permutation argument; i.e. that there are distinct isomorphic structures: '[T]here is more than one ante-rem

structure isomorphic to the natural numbers [...] An ontological realist cannot simply stipulate that there is at most one structure for each isomorphism type [...] So perhaps the realm of ante-rem structures contains both [i.e. an ante-rem structure and its permuted copy] unless simplicity or Ockham's razor can be invoked' (Shapiro 2006: 143). But as I have argued in Assadian (2018, §4), Shapiro's reply undercuts the metasemantic motivation behind the postulation of structures: If structures are, just like systems, subject to indeterminacy, then why, at least from a metasemantic perspective, postulate them in the first place? We could have lived with isomorphisms among systems and the resulting indeterminacy of reference.

- 10 See Pettigrew (2018, §4) and Leitgeb (2021, §3.1) for a similar reaction.
- 11 I owe this formulation to Richard Pettigrew. Thanks to him.
- 12 The idea that the singular terms formed on the left-hand side of an abstraction principle can be understood as making 'derived' reference to universals comes up also in Linnebo's chapter in this volume (Chapter 12), where he provides a reconciliation between Hossack's Aristotelianism on the nature of numbers as universals, and Frege's abstractionism.
- 13 This accords with Hallett's (1984: 141) interpretation of Cantor's abstractionism 'as a set theory of number'.
- 14 Thanks to John Burgess here for this formulation.
- 15 For more discussion on Finite Hume's Principle, see Heck (1997) and Mancosu (2015).

References

- Armstrong, D. M. (1986). 'In defence of structural universals'. *Australasian Journal of Philosophy* 64(1): 85–88.
- Assadian, B. (2018). 'The semantic plights of the ante-rem structuralist'. *Philosophical Studies* 175(12): 3195–3215.
- Benacerraf, P. (1965). 'What numbers could not be'. *Philosophical Review* 74(1): 47–73.
- Burgess, J. P. (1999). 'Review of Stewart Shapiro's *Philosophy of Mathematics: Structure and Ontology*'. *Notre Dame Journal of Formal Logic* 40(2): 283–291.
- Burgess, J. P. (2015). *Rigor and Structure*. Oxford: Oxford University Press.
- Cantor, G. (1955). *Contributions to the Founding of the Theory of Transfinite Numbers*. Trans. P. E. B. Jourdain. New York: Dover.
- Fine, K. (2002). *The Limits of Abstraction*. Oxford: Oxford University Press.
- Frege, G. (1953). *Foundations of Arithmetic*. Trans. J. L. Austin. Oxford: Blackwell.
- Hale, B. and Linnebo, Ø. (2020). 'Ontological categories and the problem of expressibility'. In B. Hale *Essence and Existence; Selected Essays*. Oxford: Oxford University Press, pp. 73–103.
- Hale, B. and Wright, C. (2001). *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Oxford: Clarendon Press.
- Hale, B. and Wright, C. (2009). 'The metaontology of abstraction'. In D. Chalmers, D. Manley, and R. Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Oxford University Press. pp. 178–212.
- Hale, B. and Wright, C. (2012). 'Horse sense'. *Journal of Philosophy* 109(1–2): 85–131.
- Heck, R. K. (1997). 'Finitude and Hume's principle'. *Journal of Philosophical Logic* 26(6): 589–617.
- Hellman, G. (2001). 'Three varieties of mathematical structuralism'. *Philosophia Mathematica* 9(3): 184–211.

- Hellman, G. (2005). 'Structuralism'. In S. Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press, pp. 536–562.
- Hallett, M. (1984). *Cantorian Set Theory and the Limitation of Size*. Oxford: Clarendon Press.
- Hossack, K. (2020). *Knowledge and the Philosophy of Number: What Numbers Are and How They Are Known*. London: Bloomsbury Press.
- Leitgeb, H. (2021). 'On non-eliminative structuralism. Unlabeled graphs as a case study. Part B'. *Philosophia Mathematica* 29(1): 64–87.
- Lewis, D. (1986). 'Against structural universals'. *Australasian Journal of Philosophy* 64: 25–46.
- Linnebo, Ø. (2008). 'Structuralism and the notion of dependence'. *Philosophical Quarterly* 58(230): 59–79.
- Linnebo, Ø. (2018). *Thin Objects: An Abstractionist Account*. Oxford: Oxford University Press.
- Linnebo, Ø. and Pettigrew, R. (2014). 'Two types of abstraction for structuralism'. *Philosophical Quarterly* 64: 267–283.
- Mancosu, P. (2015). 'In good company? On Hume's Principle and the assignment of numbers to infinite concepts'. *The Review of Symbolic Logic* 8(2): 1–41.
- Pettigrew, R. (2018). 'What we talk about when we talk about numbers'. *Annals of Pure and Applied Logic* 169(12): 1437–1456.
- Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press.
- Shapiro, S. (2006). 'Structure and identity'. In F. MacBride (ed.), *Identity and Modality*. Oxford: Oxford University Press, pp. 34–69.
- Shapiro, S. (2008). 'Identity, indiscernibility, and ante rem structuralism: The tale of *i* and – *i*'. *Philosophia Mathematica* 16(3): 285–309.

Arithmetic in a Finite World¹

Peter Simons

1. Kinds of quantities

Keith Hossack's ongoing work on the philosophy of mathematics takes the various kinds of number to be quantities – satisfying various but largely analogous principles – that turn on the logical features of pluralities, masses and sequences, giving rise respectively to the natural numbers, the real numbers and the ordinal numbers. Without being able to register all possible agreements and disagreements with this work, I will mention a few. I agree that natural number arithmetic comes from those properties of collections that I call *cardinalities*. A cardinality is that about a collection of things that makes it so many and neither fewer nor more. It is what enables us to make true statements about how many things are in this or that group or collection: four Beatles, twelve men who have walked on the moon, more than fifty million inhabitants of England, and so on. Standard reductionist treatments of numbers take them to be sets, which I cannot do because I don't believe in sets, and even if I did, they have the wrong properties to be or have numbers. Frege took numbers to be abstract objects, and I cannot do that either because I don't believe in abstract objects, and, even if I did, cardinalities would not be them. Frege took numbers to be objects derived from properties of concepts, and I cannot do that either because I don't believe in Frege's concepts, and, even if I did, they are not the kinds of thing of which numerical predicates other than *one* can be predicated.

Frege did sketch the beginnings of a promising treatment of real numbers as ratios of magnitudes, a view going back in its essence to Euclid, and none the worse for that. The principles such ratios have to satisfy are more complex than those for the natural numbers, and here is not the place to go into them. Good work was done early in the twentieth century not just by Frege, but also by Otto Hölder, Edward Huntington and others, and it is high time this was followed up again, as is the question of the foundations of other number systems such as the complex numbers and quaternions. This chapter will however confine itself to the counting or natural numbers, because real quantities require much more extensive treatment.

The purpose of the chapter is twofold: first, to outline an anti-realist approach to numbers deriving – surprisingly – from Frege; and second, to face (and face down) the problem that the cardinalities that support natural numbers threaten to be unavailable

if there are only finitely many individuals in the world, which was the main reason for Whitehead and Russell employing the non-logical postulate known as the Axiom of Infinity. While non-logical assumptions cannot be dispensed with completely, they can be infinitely weaker than that. By exploiting the idea of collections (pluralities) of higher order, a universe of finitely many individuals – in principle, as few as two – can furnish sufficiently many objects for all finite cardinalities to be exemplified, and so for Peano arithmetic to be guaranteed true of our world.

2. Two kinds of fiction

We are familiar with the distinction between real or actual objects, such as the Taj Mahal and John Lennon on the one hand, and fictional objects, such as Minas Tirith and Frodo Baggins, on the other. What this distinction is taken to amount to varies according to philosophical view. Some views, such as that of Roman Ingarden, hold that fictional objects exist, but in a very different manner from real objects, in that they lack the self-determination or autonomy of real objects, and depend on mental acts, of creators and readers or viewers, and texts or other physical anchors. Others, such as that of Alexius Meinong, hold that fictions do not exist, but are objects nonetheless, and are as they are independently of people, being merely thought about by people such as authors, readers etc. A third view is that fictions do not exist at all, but that people talk, write, paint, film etc. ‘about them’ in ways which are sufficiently like the way they talk etc. about real things, so that one might be forgiven for thinking they are *bona fide* objects with identities and natures. I hold this last, uncompromising view, but here is not the place to argue at length for it and against others. Rather, I shall illustrate it in use.

There are, however, fictions with different kinds of anchorage in reality. More exactly and literally, there are different ways in which our talk, thought, writing, etc. give the impression that there are objects with such and such identities and characteristics. The most fictional kinds of fiction, if we may so put it, are those corresponding to made-up stories. As an example I will use Tolkien’s Middle Earth and its places, characters and events. The brilliant and fertile brain of Tolkien brought forth stories of surpassing epic sweep, richness, detail and consistency. Despite this, they are *just stories*. There are of course similarities between the descriptions of Frodo, the Ring, rivers, trees, mountains and men, and descriptions of real people, places and things. There have to be, otherwise we would have trouble understanding them. But the descriptions do not reach out to denote anything. We cannot embark on a tour of Middle Earth to see Mount Doom and Hobbiton, or view Aragorn’s pipe in a museum. Let us call such things *mere fictions*.

Consider by way of contrast the statistical measure known as *Total Fertility Rate* (TFR). This is a way of gauging the rate at which women in a certain area, usually a country, are having children over a given period of time, usually a year.² In 2019 the highest national TFR was 7.03 in Niger and the lowest was 0.92 in South Korea. Its exact definition is not important here except to note that it is a synthetic rate applying to an ‘imaginary’ or fictitious woman (no woman can have 7.03 or 0.92 children). Artificial though it is, the TFR is based on multitudes of facts, about how

many children the women in that area are having at different ages in the year in question. So, having a basis in facts and not being simply made up, it differs from a mere fiction. To borrow a term from Leibniz, it is *well-founded*. Leibniz used the term *phenomena bene fundata* (well-founded phenomena) for bodies and other appearances such as rainbows that have their basis in his real individuals, namely monads and their characteristics. They contrast with mere appearances such as dreams and hallucinations. Because the things I am considering are not appearances, I prefer the term ‘founded fiction’.³

For purposes of the present chapter, nothing as sophisticated as TFR is being investigated, but just those very simple fictions known as the natural numbers. They are founded in the formal properties of collections of objects that we may call *finite cardinalities*.

3. Cardinality

The Beatles, the Evangelists and the constituent countries of the United Kingdom are alike in that there are four of each. They are *equinumerous*. Likewise, the months of the year, the men who walked on the Moon, and the Caesars whose biographies Suetonius wrote, are all equinumerous in being twelve. Being four, being twelve, and the like, are cardinalities, specifically, finite cardinalities. With the exception perhaps of some doctrinally subverted philosophers and mathematicians, everyone knows from their early years what it is to which cardinality properties are ascribed: it is to *collections* of things. Collections are not mathematicians’ sets, which are abstract individuals, but pluralities, of any things we like. The *set* {John, Paul, George, Ringo} is one abstract individual, but its *members*, namely John, Paul, George and Ringo, are four musicians. It is to *them*, neither to the set, nor to a Fregean concept, that the cardinality of *being four* applies. This corresponds to how we learn and use cardinal numbers, and it is in retrospect astonishing that so many philosophers have let themselves be bamboozled⁴ into denying that it is collections that have cardinalities.⁵

The cardinality of a collection is not a property of it in the way that its mass or its shape or its colour is a property of a body. To speak with the devil for a moment, it is an essential property of the collection. What this means is simply that, given this collection, it cannot but have this cardinality. The true proposition that John, Paul, George and Ringo are four has those four musicians as its truth-maker. That John is not Paul has John and Paul as its truth-maker, and so on. In being different, they are internally related, meaning simply that, given that they exist, they cannot but be different. No other items, whether properties or relations, are needed to account for those truths. This is clear, provided we specify the collection *in extension*. If we specify it through a description, we need to be more careful. That the Beatles were four⁶ is true not just because each of John, Paul, George and Ringo was a Beatle (making it true that there were *at least* four Beatles), but because there was no fifth Beatle, a proposition true by default because nothing made it false.

One of the principal objections to nominalism about numbers is that we can apply cardinality concepts to the numbers themselves: there are four primes less than ten, there are two numbers n for which the equation $a^n + b^n = c^n$ has solutions, and so on. One of the

harder jobs of a fictionalist account of arithmetic is to preserve the truth-values of these statements while eschewing platonism. For the moment, let us note only that we also commonly apply cardinalities to mere fictions as well as to founded fictions. The hobbits Frodo, Sam, Merry and Pippin, who are members of the Fellowship of the Ring, are afforded the courtesy cardinality *four*, because we understand the story that way, but there is no basis to this other than Tolkien's story. For founded fictions, it had better be better.

4. Abstraction, lite

Everyone can agree that there are *terms*, expressions, for the natural numbers: 0, 1, 2, . . . , 10, and so on. As a matter of fact, there are fewer terms (considered as tokens) than numbers, because of historical, spatiotemporal and medical limitations. But there are numerous ways of generating terms recursively so that any number could *in principle* (but not in practice) have its own term: the standard base-ten notation is just the most familiar. The important disagreement is rather about whether number terms denote, and if so, what the nature and status of their denotata are. The most promising way to furnish the numerical terms with denotata – or, as we shall hold, apparent denotata – is through abstraction. There are several things that can be called 'abstraction', but I shall employ what we may call *direct* abstraction. This goes back in its essentials to some remarks of Frege.⁷ Ironically, Frege himself abandoned it in favour of a different definition, which however proved inconsistent in his system and unwieldy in that of Whitehead and Russell. Frege's original idea has been revived in the movement in philosophy of mathematics known variously as neo-Fregeanism, neo-logicism and abstractionism.⁸

The idea – which was not in fact new with Frege – is that we can trade in a statement that an equivalence relation holds between some things for a new statement of identity between abstract objects derived from or otherwise related to the original objects and the equivalence. The nature of the trade is controversial, and I come back to it below. Consider the following examples, in semi-natural language:

The age of Pam = the age of Jim $\leftrightarrow$ Pam is as old as Jim.

The direction of line A = the direction of line B $\leftrightarrow$ line A is parallel to line B.⁹

The shape of triangle A = the shape of triangle B $\leftrightarrow$ triangle A is similar to triangle B.¹⁰

The number of the *as* = the number of the *bs* $\leftrightarrow$ there are as many *as* as there are *bs*.

The distance between A and B = the distance between C and D.

$\leftrightarrow$ A is as far from B as C is from D.

The last example is included to illustrate that the equivalences in question are not all binary relations, but can have $2n$ places for $n \geq 1$. For present purposes it is the fourth example that is interesting. Note that it differs from Frege's version (and those subsequently used by most abstractionists) in that it uses plural terms (terms for collections) rather than predicates true only of individuals.

The biconditional that I have symbolized with a double arrow is interesting. It could be just logical equivalence, but Frege put it more strongly than that. He wrote

The judgement ‘line a is parallel to line b ’, in symbols:

$$a // b$$

can be taken as an identity. If we do this, we obtain the concept of direction, and say: ‘the direction of line a is identical with the direction of line b ’. Thus we replace the symbol $//$ by the more generic symbol $=$, by distributing the specific content of the former to a and b . We carve up (*zerspalten*) the content in a way different from the original way, and gain thereby a new concept.¹¹

Frege’s metaphor of *recarving* the content is highly suggestive. The clear implication is that the two statements (judgements) have overall the same content, but that it is articulated differently. Although this predates Frege’s distinction between sense and denotation, it can only sensibly mean that the two statements have the same sense: they are synonymous. They are at the very least analytically equivalent, which may be the same thing. For this reason, I shall symbolize the recarving biconditional as ‘ $\approx$ ’ to distinguish it from mere logical equivalence.

Let us return to the statements about number. The recarving equivalence in this case has the form

$$\text{EC} \quad \#a = \#b \quad a \approx b$$

where ‘ $\#$ ’ is the ‘the number of ...’ functor and ‘ $\approx$ ’ symbolizes equinumerosity, definable as the obtaining of a bijection between the collections symbolized ‘ a ’ and ‘ b ’. EC is generally and regrettably known as Hume’s Principle;¹² for reasons of stubborn adherence to historical accuracy, I call it the *Equicardinality Principle*.¹³

Assuming the synonymy of the sides of EC, the question is whether we are to take expressions like ‘ $\#a$ ’ seriously as revealing and denoting objects that existed previously (realism), or objects introduced or somehow created by the terms in question (conceptualism), or as affording *façons de parler* which are not to be taken seriously as denoting anything (nominalism), but which are nevertheless useful. If the sides are indeed synonymous, this must mean *either* that the right-hand side of EC is, appearances to the contrary, already implicitly committed to the existence of numbers, *or* that the left-hand side, appearances to the contrary, is not. On this, people (who consider the question, of which there are probably not many) disagree. Given the extent and weight of mathematical research, the realists are always going to sit on the longer end of the lever on this, and EC is understood by abstractionists as affording an epistemological entry explaining not only how we can denote abstract objects such as (but not confined to) numbers, but also showing the link with their applications in the real world. If I know I have seventeen shirts and buy two more, I don’t need to recount them to know I now have nineteen. More needs to be said, of course, as to how predicates true of the new abstract objects get to mean what they do, but this can be managed.

Nevertheless, I find myself – perhaps foolishly, perhaps sadly – sitting on the shorter, nominalist end. Number terms, and the predicates true of them, can be treated with an ontological light touch. That $3 \neq 4$ follows quickly, that there are four prime numbers less than ten only a little more slowly. The identities and differences involved are not

mere identities and differences, but are founded in the differences among the cardinalities of collections. Contrast this with the case of mere fictions. That Frodo, Sam, Merry and Pippin are four devolves on Tolkien's creative decision. Had he chosen to extend the saga so that Sam is in fact Frodo sent back in time from the future in an Elven time-machine, then they would have been three, and only the aesthetic appraisal of Tolkien's literary wisdom would have suffered. Natural numbers cannot be messed around by arbitrary decisions: they are well-founded fictions, and it is the cardinalities of collections on which they are founded.

5. The shortage of objects problem

That being so, we have to worry about what happens if there are too few objects in the world for all the finite cardinalities to be exemplified. Suppose there are only N individuals, with N finite. Then all cardinalities greater than N are unexemplified. There is no a with sufficiently many members for $\#a$ to be greater than N . Then Peano arithmetic is falsified, for there is no successor to the number $N = \#U$, where U is the collection of all individuals. It was for this reason that in *Principia Mathematica* Whitehead and Russell postulated an Axiom of Infinity, to the effect that every natural number is the number of some class, and the natural numbers are all distinct so there are infinitely many individuals. This assumption was widely judged, not least by Russell himself, to be non-logical and *a posteriori*.¹⁴

Frege had secured an infinity of objects by a different route, through his value ranges. These delivered objects for functions, including concepts, and so could provide an object for the empty concept, and commence an unending bootstrap generating ever larger classes, including the class of all finite numbers. This has what Russell would call an air of hocus-pocus about it: an infinite number of individuals are extracted from none at all. That arithmetic may follow on abstractionist lines from EC rather than from Frege's inconsistent logic, a result that George Boolos calls¹⁵ *Frege's Theorem*, does not detract from this air, since it is precisely the status of the terms ' $\#a$ ' etc. on the left-hand side of EC that is questionable. Are they genuinely denoting singular terms or not? Does the truth of a statement of the form ' $\#a = \#b$ ' depend on what the terms denote, as it does in cases like 'Eric Blair = George Orwell', or does it inherit its truth from elsewhere – like, for instance, the right-hand side of EC?

In the absence of a (neo-)Fregean bootstrap, the Whitehead–Russell reliance on the Axiom of Infinity is more honest, but has the disturbing consequence that Peano arithmetic may be empirically false, if the world is finite in individuals.

6. Zero, one, many, manys, manyses

I started by saying that cardinalities are properties of collections, and not of concepts or sets. What about the case of one and zero, of which Frege was so proud that his theory was better than that of, say, Husserl? Surely a single individual is not a collection? Not in the ordinary sense of the word: it is an exception. But every individual has its

number: one. So, with a little stretch of terminology, we can call an individual a collection of one. Then there is the number zero. There are certainly no collections, even in the extended sense, with no members. But that does not stop the operator ‘#’ from making sense. Suppose a are the round squares and b are the golden mountains. Then $\#a = \#b = 0$, because there are exactly as many round squares as there are golden mountains, and as there are non-self-identical objects, namely none at all. One and zero are not problematic for abstractionism lite.

Some books are written by single authors, some are jointly authored. Frege wrote *Begriffsschrift* alone, whereas Whitehead and Russell wrote *Principia Mathematica* together. What about the collection of joint authors of logic books? Apart from Whitehead and Russell, they include, among others, Arnauld and Nicole, Lewis and Langford, Hilbert and Bernays, Rasiowa and Sikorski, Hilbert and Ackermann, not to mention the five members of L. T. F. Gamut. Notice that Hilbert occurs more than once. Suppose we consider all these collections of joint authors. That is then a collection of second order, a collection of collections. There will be a similar collection of joint authors of physics books. That and the collection of joint authors of logic books makes a collection of collections of collections, a third-order collection. And so on. The reasons to accept that collections of individuals are distinct from individuals are the same reasons that force us to accept that second-order collections are different from first-order ones. Difference drives everything. The joint author pair Hilbert–Ackermann is different from the joint author pair Hilbert–Bernays because they differ in at least one member. And the logic joint authors collection has different members from the physics joint authors collection for the same reason, though there are few or no overlaps here. Collections (of more than one, i.e. pluralities) are different objects from individuals. In that sense, they are new and additional entities. But they are not *wholly* new, and their existence is not optional. Given the individual members, the collection exists automatically. Even God could not create Whitehead and create Russell but fail to create Whitehead and Russell. If the two individuals exist, so does the pair.

Considering Whitehead and Russell, let’s use the initials instead of the names. The pair who wrote *Principia Mathematica* we write as WR. Once we accept that difference engenders collections, there is no limit to the order of collections. There only need to be enough individuals to start with. Since $W \neq R$, $W \neq WR$ and $R \neq WR$, so we have three items: Whitehead, Russell, and the pair. They are all different from one another, despite the overlaps. So we have $W R W R$ as a second-order collection of three items, despite the overlaps. This can be continued: $WR \neq W W R$, so $W W R | W R$ is a third-order collection. And so on. In orders up to and including 3, there are 59 collections based on just those two gentlemen, and the number rises rapidly at each new order. Each such order, based on two or any finite number of individuals, is finite, but since there is no upper limit to the orders, it means two or more individuals lie at the basis of infinitely many collections.

7. The axiom of two

Whitehead and Russell needed infinitely many individuals to guarantee the truth of Peano arithmetic. We can pare that number back to two individuals, provided we

accept the open hierarchy of orders of collections. One individual is impotent to generate such a hierarchy however, and no individuals more obviously so. There may be metaphysical reasons why there have to be two or more individuals, but there are no logical reasons. A one-element universe or even an empty universe is a logically consistent possibility. Therefore, on our account, Peano arithmetic is not necessarily true, but requires as an additional premiss that there are at least two individuals. Given that the actual world contains many more than two individuals, Peano arithmetic is safely true of our world, even if that world is finite in its individuals.

8. Back to kindergarten

In a late note, 'Zahlen und Arithmetik',¹⁶ probably written in the last year of his life, Frege wrote disparagingly about the natural numbers, on whose logical derivation he had spent (fruitless, he thought) decades. He called them 'die Kleinkinder-Zahlen', literally, the little-children-numbers. The English translators render this rather well as *kindergarten-numbers*.¹⁷ These, he says, are drilled into¹⁸ little children by parents and teachers, in order to prepare them for a future career, most likely in commerce. He makes unfair and unsympathetic fun of how small children are taught to count and compute.

Leaving Frege's late nastiness aside, what on earth is wrong with kindergarten and how (natural number) arithmetic is taught? Nothing, of course. You don't have to read the ontology off the pedagogy, but they shouldn't lose touch. To add, you illustrate with, and need, two disjoint collections. To multiply, you need several disjoint equinumerous collections: that's order two. To illustrate associativity of multiplication, you need order three. The pedagogy has to work to drill the principles so that they conform to the algebra of numerical operations and can be applied to concrete cases. It seems to work perfectly well. Kindergarten arithmetic however does not usually, at least initially, inculcate in children the belief that there are infinitely many natural numbers. There is a telling exchange about this in the Disputation introducing Heyting's *Intuitionism*:¹⁹

Int Children in the elementary school understand what the natural numbers are and they accept the fact that the sequence of natural numbers can be indefinitely continued.

Letter It is suggested to them that they understand.

Perhaps the finitist LETTER would consider that this belief is *eingeschmiert*, rather than rationally justified.

9. An infinite sequence from two

Let's use the Sun (symbol ☉) and the Moon (symbol ☾) as two individuals to start things off. For an orderly sequence of objects, we start with the Sun, then add the Moon

to give a pair, then for each subsequent collection take all the objects mentioned in their predecessors. So the sequence – call it A – starts

$\odot$
 $\odot \subset$
 $\odot \mid \subset \mid \odot \subset$
 $\odot \parallel \subset \parallel \odot \subset \parallel \odot \mid \subset \mid \odot \subset$
 $\odot \parallel \parallel \subset \parallel \parallel \odot \subset \parallel \parallel \odot \mid \subset \mid \odot \subset \parallel \parallel \odot \mid \subset \mid \odot \subset$

and so on.

The vertical dividers punctuate like the groups of dots in Whitehead and Russell. The cardinalities of these collections are one to five respectively, and in their recursive embeddedness they resemble the von Neumann numbers. We have the following truths: $\text{One}(\odot)$, $\text{One}(\subset)$, $\text{Two}(\odot \subset)$, $\text{Three}(\odot \mid \subset \mid \odot \subset)$, $\text{Four}(\odot \parallel \subset \parallel \odot \subset \parallel \odot \mid \subset \mid \odot \subset)$, and so on. Letting a be any collection in the sequence, and $\preceq(a)$ the collection of it and its predecessors, if $n(a)$ then $S(n)(\preceq(a))$, where $S(n)$ is the successor cardinality to n . If Λ is the ‘collection’ of all non-self-identical objects, then $\text{Zero}(\Lambda)$, and we stipulate that $\text{One} = S(\text{Zero})$. Since it is evident that for any $a \in A$, $a \neq \preceq(a)$ and $\preceq(a)$ has one more member than a , the whole sequence A is infinite and models the natural numbers.²⁰ In a world with two or more individuals, every finite cardinality is exemplified.

Any collection c has finite cardinality n if and only if for some member a of A , $c \approx a$ and $n(a)$. But A is not the sequence of natural numbers. Even speaking liberally and tolerating founded fictions, A is still not the sequence of natural numbers, and nor are the cardinalities of the members of A , because we are missing the abstraction step. For this, we need EC.

10. From Frege to Peano

We recall

$$\text{EC} \quad \#a = \#b \quad a \approx b$$

For the numbers as fictions, we use the standard numerals, and for the number variable corresponding to the cardinality predicate variable, we use upper case. We have

$$\#a = 0 \leftrightarrow \text{Zero}(a)$$

$$\#a = 1 \leftrightarrow \text{One}(a)$$

$$\#a = 2 \leftrightarrow \text{Two}(a)$$

etc., and in general

$$\#a = N \leftrightarrow n(a)$$

$$\#a = \sigma(N) \leftrightarrow \neg b \preceq a \wedge n(b) \wedge a \approx \preceq(b)$$

where σ is the successor function on the numbers.²¹

To ‘bootstrap’ the numbers out of nowhere in Frege’s style, we follow Boolos’s formulation and define zero, the relation ‘immediately precedes’ and ‘natural number’:

Def. 0 $0 = \#(\Lambda)$

Def. P $\forall x \forall y \ulcorner xPy \leftrightarrow$
 $a \ b \ z \urcorner z \varepsilon b \wedge \forall w \ulcorner w \varepsilon a \leftrightarrow w \varepsilon b \wedge w \neq z \urcorner \wedge x = \#a \wedge y = \#b \urcorner$

Def. N $\forall x \ulcorner x \varepsilon N \leftrightarrow \forall a \ulcorner 0 \varepsilon a \wedge \forall y \forall z \ulcorner y \varepsilon a \wedge yPz \rightarrow z \varepsilon a \urcorner \rightarrow x \varepsilon a \urcorner$

The definitions of P and N quantify over collections, and so belong to the plurals equivalent of second-order logic. The latter definition employs the standard definition of the ancestral.

What Boolos calls ‘Frege’s Theorem’ is that the Peano axioms follow from second-order logic, EC, and these definitions.²² The fictions that are the natural numbers are well-founded provided there are at least two individuals, and provided higher-order collections are admitted. Otherwise, without higher-order collections, we are back to the Axiom of Infinity, and the natural numbers are then not well-founded in any finite universe.

Notes

- 1 To four: Jacob, Wolfgang, Torben and Freya.
- 2 It differs from the more straightforward *birth rate*, which divides the number of live births in a population by the number of the population by the number of years.
- 3 Leibniz did also use the term *fictiones bene fundata*, but more restrictedly, for useful mathematical fictions like infinity and infinitesimals.
- 4 Principally, I think, by Frege. Early Russell had a distinction between classes as one and classes as many: ‘A class [...] is distinct from the whole composed of its terms, for the latter is only and essentially one, while the former, where it has many terms, is [...] the very kind of object of which *many* is to be asserted’ (Russell 1903: 680). Even while writing the *Principles*, Russell converted to Frege’s view.
- 5 An honourable exception is Husserl (1970: 15): ‘As to the concrete phenomena that form the basis for the abstraction of the concepts in question [sc. the cardinal numbers – PS] there is no doubt. They are collections [*Inbegriffe*], pluralities [*Vielheiten*] of certain objects. Everyone knows what is meant by this expression.’ Of course everyone – among analytic philosophers – ‘knows’ that Husserl was refuted by Frege, and indeed there is too much psychology in Husserl’s account, but on this crucial point he is spot on.
- 6 Throughout their definitive (1962–70) line-up.
- 7 Frege (1953) §§62–66.
- 8 The *fons et origo* of this is Wright (1983).
- 9 Frege’s example: *ibid.*, § 64. He is clearly here ignoring the sense (‘direction’) of a directed line, which is involved in what we *normally* consider direction (north is not south but any two N–S lines are parallel).

- 10 Ibid.
- 11 Ibid., §64, pp. 74 f. The translation is a slight modification of Austin's.
- 12 The misleading name stems from Boolos, but the idea goes back to Frege again (ibid., § 62, p. 75), who misinterprets a passage from Hume. Hume is talking about comparing *numbers*, not classes.
- 13 If anyone's name deserves to be attached to it, it is that of Cantor: cf. his (1932): 283.
- 14 Russell (1919): 141–143.
- 15 Boolos (1996): 277. See also Boolos (1995).
- 16 Frege (1983).
- 17 Frege (1979): 276.
- 18 Frege's own term here is *eingeschmiert* (Frege 1983: 296) – literally, smeared (or rubbed) in – which is decidedly more pejorative than 'drilled', and I can't find a similarly pejorative English translation. 'Drilled' is along the right lines: think of sergeant majors, with a hint of arbitrariness thrown in. Very unbalanced from Frege.
- 19 Heyting (1956): 7.
- 20 The constant 'ε' stands for the singular inclusion predicate: ' $x \varepsilon a$ ' means that x is one of (is a member of) a .
- 21 The corners mark quantifier scope.
- 22 The proof is first outlined in Wright (1983): 154 ff.

References

- Boolos, G. (1995). Frege's Theorem and the Peano Postulates. *Bulletin of Symbolic Logic* 1: 317–326. Reprinted in *Logic, Logic, and Logic*. Cambridge: Harvard University Press, 1998, pp. 291–300.
- Boolos, G. (1998). 'On the proof of Frege's theorem'. Reprinted in *Logic, Logic, and Logic*. Cambridge: Harvard University Press, 1998, pp. 275–291.
- Cantor G. (1932). 'Beiträge zur Begründung der transfiniten Mengenlehre'. In E. Zermelo (ed.), *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Berlin: Springer, pp. 282–356.
- Frege, G. (1953). *The Foundations of Arithmetic*, 2nd ed., trans. J. L. Austin. Oxford: Blackwell.
- Frege, G. (1979). *Posthumous Writings*. Oxford: Blackwell.
- Frege, G. (1983). *Nachgelassene Schriften*, 2nd ed. Hamburg: Meiner.
- Heyting, A. (1956). *Intuitionism*. Amsterdam: North-Holland.
- Husserl, E. (1970). *Philosophie der Arithmetik*, ed. L. Eley. The Hague: Nijhoff.
- Russell, B. (1903). *The Principles of Mathematics*. London: Allen & Unwin.
- Russell, B. (1919). *Introduction to Mathematical Philosophy*. London: Allen & Unwin.
- Wright, C. (1983). *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press.

Index

- Aardvark 53, 58
Aczel, Peter 91
Adams, Robert M. 122, 139
Annas, Julia 63, 155, 157
Aristotle 4, 7, 10, 34, 63, 148, 189–190, 192, 199
Armstrong, David M. 87, 90, 169–171
Aquinas, Thomas 87
Ayer, Alfred J. 171–172
- Benacerraf, Paul 25, 207
Bennett, Jonathan 137
Benson, Hugh 156
Bergmann, Gustav 173
Boghossian, Paul 63
Boolos, Georg 93, 98, 177, 179, 222, 226
Bradley, Francis H. 6, 34
Brentano, Franz 13–15, 31, 87–98
Bucephalus 163, 165
Burali-Forti, Cesare 24, 198, 199
Burgess, John 25–26, 204, 211–212
Burnyeat, Miles 64, 146, 156, 157, 158
- Cantor, Georg 92–93, 212
Cassio 16, 102–103, 114
Cleve, James van 91
Coco 95
Craig, Leon H. 158
Cross, Robert C. 155, 157, 158
- Davidson, Donald 15, 39
Dedekind, Richard 27, 194, 207
Descartes, René 7
Desdemona 16, 102–103, 114
Donnellan, Keith 115
Dummett, Michael 115, 139
- Eleatic Stranger 54, 162
Elkind, Landon D.C. 172
- Euclid 10–11, 139, 143, 146–147, 152, 154, 156, 159, 191
Evans, Gareth 38, 107
- Fine, Gail 62, 64, 155, 156, 157
Fine, Kit 119, 134, 193, 205, 213
Fraenkel, Abraham 8–9, 15
Frege, Gottlob 8, 10–11, 16–17, 23–26, 106–107, 125, 139, 165, 190, 195, 198, 205, 217, 220–222, 224
- Giaquinto, Marcus 157, 159
Glaucón 56, 63, 65, 148, 150
Gödel, Kurt 19
Goldbach, Christian 19, 139–140
Goodman, Nelson 18–19, 118–121, 123–124, 126, 128–129, 131–137, 141
Grube, Georg 64
- Hale, Bob 25, 205, 209, 213
Hamilton, William 87–88
Harte, Verity 64, 149, 153, 157
Hellman, Geoffrey 25, 204, 207–208
Hempel, Carl 117–118
Heyting, Arend 224
Hochberg, Hans 169–171, 173
Hume, David 23–24, 26, 117–118, 124–125, 127
Husserl, Edmund 222
- Incitatus 163, 165
- Jamina 12
Johnston, Mark 42–43
Juliet 6
Julius 107–113, 115
- Kant, Immanuel 17, 106–107
Kim, Jaegwon 90–91
Kitcher, Philip 108–110

- Korsgard, Christine 62
 Kripke, Saul 9, 17–18, 22, 179
- Leibniz, Gottfried W. 15, 17, 106–107, 122, 124–125
 Leśniewski, Stanisław 10
 Lewis, David 10, 18, 119–121, 123–124, 128, 133–139
 Lillehammer, Hallvard 1
 Locke, John 87
- Marcus, Ruth Barcan 27, 187, 188
 Meletus 2–3
 Meno 51–53
 Mill, John Stuart 88, 117, 124–125, 138
 Moore, George E. 33
 Müller–Lyer, Franz C. 42
- Netz, Reviel 146–147, 157
 Nixon, Richard 134
- Patterson, R. 146–147, 155, 156, 157, 159
 Pettigrew, Richard 205, 211
 Plato 2, 4–5, 7, 10, 12, 20, 49–65, 143–159, 161, 166, 101
Apology of Socrates, The 65
Charmides 63
Euthydemus 58–61, 146, 152
Gorgias 63, 65
Laws 157
Meno 5, 51–53, 57, 61–65, 147, 151–154, 155, 157, 159
Parmenides 150
Phaedo 62, 64, 151–152, 154, 158
Phaedrus 4, 166
Philebus 145, 156, 157
Protagoras 63, 65
Republic 12, 20, 49–51, 55–65, 143–159
Seventh Letter 153
Sophist 54, 161
Symposium 58
Theaetetus 5, 51–54, 56–58, 60–61, 63–65, 157
Timaeus 155, 157
 Pharaoh 38, 43
 Plantinga, Alvin 121–122, 128
- Prior, Arthur N. 28
 Pritchard, Paul 62, 64, 157, 158
 Proclus 152
 Prodicus 159
- Quine, Willard v. O. 10, 24, 27, 141, 199
- Ramsey, Frank 117
 Reid, Thomas 7, 13–14, 87–88
 Romeo 6
 Rosen, Gideon 122
 Rumfitt, Ian 22
 Russell, Bertrand 1–3, 5–7, 16, 21, 24, 26–27, 34, 101–103, 114, 161–173, 220, 222–223
- Shaprio, Stewart 25, 203–204, 208–209, 211
 Simmias 150, 152
 Skittles 95
 Socrates 2–4, 7, 10, 20, 49–65, 101–104, 143–159, 166, 173
 Stalnaker, Robert 119–122, 133, 138–139
 Stumpf, Carl 95
- Tarski, Alfred 10, 19, 34
 Theaetetus 2–3, 53, 54, 158, 162
 Thrasy machus 2–3, 57
- Ueberweg, Friedrich 96
 Uzquiano, Gabriel 22
- von Neumann, John 206
- Whitehead, Alfred N. 21, 220, 222–223
 Williams, Bernard 62
 Williamson, Timothy 22, 39, 125, 138, 177–178, 182, 185
 Wittgenstein, Ludwig 21, 114, 161–162
 Woozley, Anthony D. 155, 157, 158
 Wright, Crispin 25, 205, 209, 213
- Zagzebski, Linda 62
 Zermelo, Ernst 8–9, 15, 204, 206

